

UNIVERSIDADE FEDERAL DO PARANÁ

RAFAEL ROMUALDO WANDRESEN

*SOFTWARE* EDUCACIONAL PARA O ENSINO DE INVENTÁRIO FLORESTAL

CURITIBA  
2014

RAFAEL ROMUALDO WANDRESEN

*SOFTWARE EDUCACIONAL PARA O ENSINO DE INVENTÁRIO FLORESTAL*

Tese apresentada como requisito parcial à obtenção do grau de Doutor em Engenharia Florestal, Programa de Pós-Graduação em Engenharia Florestal, Setor de Ciências Agrárias, Universidade Federal do Paraná.

Orientador:

Prof. Dr. Henrique Soares Koehler

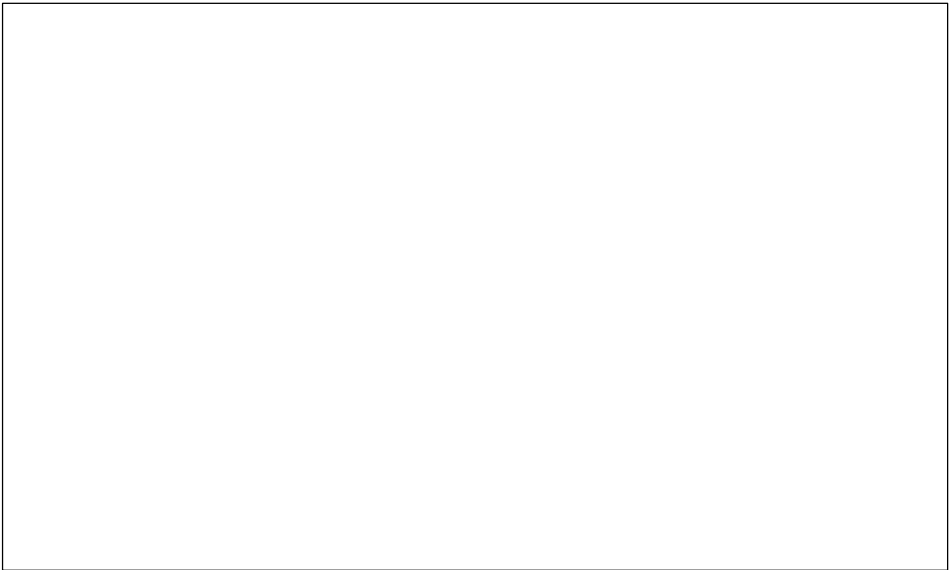
Coorientadores:

Prof. Dr. Sylvio Péllico Netto

Prof. Dr. Carlos Roberto Sanquetta

CURITIBA  
2014

Ficha catalográfica elaborada por XXXX – CRB XXXX/PR



**PARECER**

Dedico este trabalho  
aos meus filhos Maria Fernanda e João Henrique  
à minha esposa Viviane  
e aos meus pais Romualdo e Maria Otília.

## AGRADECIMENTOS

Agradeço à Universidade Federal do Paraná e ao programa de pós-graduação em Engenharia Florestal pela estrutura física e corpo docente comprometidos com a formação dos alunos.

Ao professor Dr. Henrique Koehler pela orientação precisa e objetiva, e, principalmente por acreditar e defender o meu trabalho.

Ao professor Dr. Sylvio Péllico Netto pela disposição no compartilhamento da sua vasta experiência na área florestal e por todas as horas investidas para me ajudar a resolver os desafios desse trabalho.

Ao professor Dr. Carlos Roberto Sanquetta pela sua coorientação e pelas observações valiosas ao *software* desenvolvido.

À professora Dra. Ana Paula Dalla Corte pela sua disposição em testar o *software* desenvolvido com uma turma da graduação.

Ao professor Dr. Sebastião do Amaral Machado por permitir a apresentação do *software* em uma turma da pós-graduação.

Ao secretário do curso de pós-graduação em Engenharia Florestal, Reinaldo Pinheiro, pelos atendimentos às solicitações burocráticas de forma rápida e eficaz.

Quero expressar meu profundo agradecimento à minha família pelo apoio incondicional, pela compreensão das ausências e pelo carinho em todos os momentos. Com muito carinho e amor, agradeço à minha esposa Viviane; aos meus dois filhos queridos, João Henrique e Maria Fernanda; e aos meus pais Romualdo e Maria Otília.

## RESUMO

Os processos de amostragem são um importante conteúdo na disciplina de Inventário Florestal dos cursos de Engenharia Florestal. O principal objetivo deste trabalho foi desenvolver e avaliar um *software* educacional para a disciplina de Inventário Florestal, como um recurso didático no qual os alunos de graduação e pós-graduação possam exercitar 16 processos de amostragem. A aplicação desenvolvida não precisa ser instalada e é acessada pela internet. Os dados utilizados são gerados individualmente para cada aluno por meio de geração pseudoaleatória de dados, valendo-se do método “núcleo estimador variável” e a correção dos exercícios é realizada de forma automática. Uma biblioteca de programação para os processos de inventário florestal foi modelada e desenvolvida considerando-se os conceitos de orientação a objetos. O modelo APDIA (Análise, Projeto, Desenvolvimento, Implementação e Avaliação) e a técnica de prototipagem da Engenharia de *Software* foram utilizados na construção dessa aplicação. Utilizou-se a pesquisa avaliativa, por meio de um instrumento de avaliação que foi aplicado aos alunos da disciplina de Inventário Florestal, do curso de graduação em Engenharia Florestal da Universidade Federal do Paraná. Essa avaliação revelou que a percepção dos alunos foi positiva em relação ao *software* nos seguintes critérios: benefício educacional, facilidade de uso e interesse futuro. O sistema de pontuação do *software* teve a concordância parcial entre os alunos.

Palavras-chave: Inventário Florestal. *Software* Educacional. Objeto de Aprendizagem. Processos de Amostragem

## **ABSTRACT**

Sampling procedures are important in the scope of the Forest Inventory course in forestry. The main goal of this work was to develop and to evaluate a learning software for a Forest Inventory course, as a teaching resource in which the undergraduate and graduate students could work with 16 sampling procedures. This study reports the implementation and the use of a learning software, which offers to the student 16 exercises about the sampling techniques. The developed application does not need to be installed and it is accessed through the internet. The used data are individually generated for each student through pseudorandom generation data using the method "variable kernel density estimation". The correction of the exercises is performed automatically. A programming library for the forest inventory procedures was modeled and developed considering the concepts of object orientation. The ADDIE (Analysis, Design, Development, Implementation and Evaluation) model and the prototyping technique of software engineering were used in the construction of this software. An evaluative research was used during the test application, effectuated with students of the Forest Inventory course at the School of Forestry of the Federal University of Paraná. This survey confirmed that the software provoked a positive perception by the students regarding educational benefits, ease of use and appropriate for learning. According to the students' opinion, the scoring system was only partially accepted as a good tool for evaluation.

Kew-words: Forest inventory. Educational Software. Learning object. Sampling Procedures.



## LISTA DE FIGURAS

FIGURA 1 – ALGORITMO PARA CÁLCULO DA INTENSIDADE DE AMOSTRAGEM .....	36
FIGURA 2 – MODELO EM CASCATA DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE .....	63
FIGURA 3 – MODELO INCREMENTAL DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE .....	64
FIGURA 4 – MODELO DE PROTOTIPAÇÃO DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE .....	66
FIGURA 5 – MODELO ESPIRAL DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE .....	67
FIGURA 6 – PROCESSO UNIFICADO PARA CONSTRUÇÃO DE SOFTWARE.....	68
FIGURA 7 – EXEMPLO DE CLASSE QUE REPRESENTA UMA ÁRVORE .....	74
FIGURA 8 – EXEMPLO DE CLASSE QUE REPRESENTA UM RETÂNGULO.....	75
FIGURA 9 – EXEMPLO DE RELACIONAMENTO DE DEPENDÊNCIA ENTRE CLASSES.....	78
FIGURA 10 – EXEMPLOS DE ASSOCIAÇÃO SIMPLES ENTRE CLASSES.....	79
FIGURA 11 – EXEMPLO DE AGREGAÇÃO ENTRE CLASSES.....	80
FIGURA 12 – EXEMPLO DE COMPOSIÇÃO ENTRE CLASSES .....	81
FIGURA 13 – EXEMPLO DE GENERALIZAÇÃO ENTRE CLASSES.....	82
FIGURA 14 – EXEMPLO DE CLASSE ABSTRATA.....	84
FIGURA 15 – EXEMPLO DE CLASSE ANINHADA .....	85
FIGURA 16 – EXEMPLO CLASSE GENÉRICA.....	86
FIGURA 17 – TEOREMA DO MÉTODO DA INVERSA DA PROBABILIDADE CUMULATIVA .....	90
FIGURA 18 – EXEMPLO DE GRÁFICO DO MÉTODO DA INVERSA DA FDA.....	90
FIGURA 19 – COMANDO NO <i>SOFTWARE R</i> PARA GERAÇÃO DE NÚMEROS PSEUDOALEATÓRIOS DE UMA DISTRIBUIÇÃO NORMAL.....	91
FIGURA 20 – FLUXOGRAMA DAS ETAPAS DO TRABALHO SEGUNDO A METODOLOGIA ADOTADA .....	99

FIGURA 21 – DIAGRAMA SIMPLIFICADO DE COMPONENTES DO SOFTWARE .....	106
FIGURA 22 – PASSOS IMPLEMENTADOS NO SOFTWARE EDUCACIONAL PARA GERAÇÃO DE DADOS PSEUDOALEATÓRIOS.....	108
FIGURA 23 – CLASSE QUE REPRESENTA A UNIDADE AMOSTRAL.....	113
FIGURA 24 – RELAÇÃO ENTRE AS CLASSES QUE REPRESENTAM A UNIDADE AMOSTRAL E A UNIDADE AMOSTRAL DE ÁREA FIXA .....	114
FIGURA 25 – RELAÇÃO ENTRE AS CLASSES QUE REPRESENTAM A AMOSTRA E SUAS UNIDADES AMOSTRAIS .....	116
FIGURA 26 – CLASSE QUE REPRESENTA A AMOSTRAGEM SIMPLES E SUAS RELAÇÕES.....	118
FIGURA 27 – CLASSE QUE REPRESENTA A POPULAÇÃO .....	119
FIGURA 28 – CLASSES QUE REPRESENTAM OS ESTRATOS E SUAS RELAÇÕES.....	119
FIGURA 29 – CLASSE QUE REPRESENTA A AMOSTRAGEM ESTRATIFICADA .....	121
FIGURA 30 – CLASSES QUE REPRESENTAM A POPULAÇÃO POR FAIXA, A POPULAÇÃO POR UNIDADE PRIMÁRIA E A POPULAÇÃO POR IDADE .....	122
FIGURA 31 – CLASSES QUE REPRESENTAM A AMOSTRAGEM ESTRATIFICADA E A AMOSTRAGEM ESTRATIFICADA DE TAMANHOS DIFERENTES.....	123
FIGURA 32 – CLASSES QUE REPRESENTAM A AMSTRAGEM EM DOIS ESTÁGIOS, A AMOSTRAGEM EM CONGLOMERADOS E A AMOSTRAGEM SISTEMÁTICA COM MÚLTIPLOS INÍCIOS ALETÓRIOS.....	124
FIGURA 33 – CLASSES QUE REPRESENTAM A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES .....	125
FIGURA 34 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES INDEPENDENTES .....	126
FIGURA 35 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO TOTAL .....	127
FIGURA 36 – CLASSE QUE REPRESENTA A AMOSTRAGEM DUPLA EM MÚLTIPLAS OCASIÕES.....	128
FIGURA 37 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO PARCIAL .....	129

FIGURA 38 – HISTOGRAMA DOS DADOS ORIGINAIS (ANEXO 1) ESTRATIFICADOS POR IDADE .....	132
FIGURA 39 – COMPARATIVO DOS DADOS ORIGINAIS (ANEXO 1) E DADOS GERADOS (APÊNDICE 6) POR MEIO DA FUNÇÃO ACUMULADA DE PROBABILIDADE .....	134
FIGURA 40 – ESQUEMA BÁSICO DE FUNCIONAMENTO .....	135
FIGURA 41 – FLUXOGRAMA DE FUNCIONAMENTO .....	135
FIGURA 42 – (A) TELA DE LOGIN (B) TELA DE CRIAÇÃO DE NOVO USUÁRIO	137
FIGURA 43 – TELA INICIAL .....	137
FIGURA 44 – PASSO 1 DO DESAFIO 2 (PARÂMETROS DA POPULAÇÃO ESTRATIFICADA).....	139
FIGURA 45 – PASSO 2 DO DESAFIO 2 (PARÂMETROS DA POPULAÇÃO ESTRATIFICADA).....	140
FIGURA 46 – PASSO 1 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES) .....	142
FIGURA 47 – PASSO 2 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES) .....	142
FIGURA 48 – (A) PASSO 3 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES) (B) MENSAGEM DO SISTEMA.....	143
FIGURA 49 – PASSO 4 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES) .....	143
FIGURA 50 – PASSO 1 DO DESAFIO 10 (AMOSTRAGEM EM DOIS ESTÁGIOS) .....	144
FIGURA 51 – PASSO 2 DO DESAFIO 10 (AMOSTRAGEM EM DOIS ESTÁGIOS) .....	145
FIGURA 52 – PASSO 1 DO DESAFIO 14 (AMOSTRAGEM EM MÚLTIPLAS OCASIÕES – AMOSTRAGEM COM REPETIÇÃO TOTAL) .....	146
FIGURA 53 – PASSO 2 DO DESAFIO 14 (AMOSTRAGEM EM MÚLTIPLAS OCASIÕES – AMOSTRAGEM COM REPETIÇÃO TOTAL) .....	147
FIGURA 54 – TELA DE TABELA COMPARATIVA DA PONTUAÇÃO.....	147
FIGURA 55 – EXEMPLO DO USO DE NÚCLEO ESTIMADORES VARIÁVEL .....	165
FIGURA 56 – PROBABILIDADE ENTRE 30 E 35 CM. ....	168

FIGURA 57 – DIAGRAMA DE ENTIDADE-RELACIONAMENTO – MODELO FÍSICO	
.....	170
FIGURA 58 – EVIDÊNCIA DOS TESTES UNITÁRIOS COM ERROS E FALHAS .	174
FIGURA 59 – EVIDÊNCIA DOS TESTES UNITÁRIOS SEM ERROS OU FALHAS	
.....	175

## LISTA DE TABELAS

TABELA 1 – EXEMPLO DO TESTE DE KOLMOGOROV-SMIRNOV PARA OS DADOS GERADOS.....	133
TABELA 2 – RESULTADOS DA AVALIAÇÃO (N=23) .....	149
TABELA 3 – EXEMPLO DE CÁLCULO DO NÚCLEO ESTIMADOR VARIÁVEL....	167
TABELA 4 – EXEMPLO DE GERAÇÃO DE NÚMEROS PSEUDOALETÓRIOS COM BASE NO MÉTODO NÚCLEO ESTIMADOR VARIÁVEL .....	169
TABELA 5 – RESULTADO DO TESTE DE KOLMOGOROV-SMIRNOV PARA 102 CONJUNTOS DE DADOS GERADOS .....	178

## LISTA DE QUADROS

QUADRO 1 – PARÂMETROS DA POPULAÇÃO POR UNIDADE AMOSTRAL .....	27
QUADRO 2 – PARÂMETROS DA POPULAÇÃO ESTRATIFICADA .....	28
QUADRO 3 – PARÂMETROS DA POPULAÇÃO POR FAIXA .....	29
QUADRO 4 – PARÂMETROS DA POPULAÇÃO POR FAIXA COM TAMANHOS DIFERENTES .....	30
QUADRO 5 – EXEMPLO DE UMA POPULAÇÃO COM UNIDADES PRIMÁRIAS E SECUNDÁRIAS .....	30
QUADRO 6 – PARÂMETROS DA POPULAÇÃO POR UNIDADE PRIMÁRIA .....	32
QUADRO 7 – ESTIMATIVAS PARA AMOSTRAGEM ALEATÓRIA SIMPLES .....	35
QUADRO 8 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM ALEATÓRIA SIMPLES.....	36
QUADRO 9 – ESTIMATIVAS PARA AMOSTRAGEM ESTRATIFICADA .....	39
QUADRO 10 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM ESTRATIFICADA.....	40
QUADRO 11 – ESTIMATIVAS PARA AMOSTRAGEM SISTEMÁTICA.....	43
QUADRO 12 – ESTIMATIVAS PARA AMOSTRAGEM EM DOIS ESTÁGIOS .....	46
QUADRO 13 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM DUPLA .....	46
QUADRO 14 – ESTIMATIVAS PARA AMOSTRAGEM EM CONGLOMERADOS....	48
QUADRO 15 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM EM CONGLOMERADOS.....	49
QUADRO 16 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES INDEPENDENTES.....	53
QUADRO 17 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO TOTAL.....	55
QUADRO 18 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM AMOSTRAGEM DUPLA .....	58
QUADRO 19 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO PARCIAL.....	61

QUADRO 20 – ESPECIFICADORES DE ACESSO DE MEMBROS DE CLASSE ...	76
QUADRO 21 – TIPOS DE DADOS PRIMITIVOS NA LINGUAGEM JAVA .....	77
QUADRO 22 – INSTRUMENTO DE AVALIAÇÃO .....	103
QUADRO 23 – DADOS DE ORIGEM UTILIZADOS DO LIVRO TEXTO (PÉLLICO NETTO E BRENA, 1997, P. 97-303) POR EXERCÍCIO IMPLEMENTADO.....	107
QUADRO 24 – PROPRIEDADES DA ESTRUTURA DA POPULAÇÃO QUE SÃO SORTEADOS.....	109
QUADRO 25 – PROPRIEDADES DA ESTRUTURA DA POPULAÇÃO PARA O “ALUNO 1”.....	131
QUADRO 26 – DESAFIOS E QUESTÕES .....	136
QUADRO 27 – ESTADOS DOS DESAFIOS.....	138
QUADRO 28 – ORGANIZAÇÃO DOS PROCESSOS DE AMOSTRAGEM .....	138
QUADRO 29 – CLASSES IMPLEMENTADAS NA BIBLIOTECA DE PROGRAMAÇÃO COM UMA BREVE DESCRIÇÃO .....	172
QUADRO 30 – EVIDÊNCIA DOS TESTES REALIZADOS NA BIBLIOTECA DE PROGRAMAÇÃO QUANDO EXECUTADA PARA RESOLVER OS EXERCÍCIOS PROPOSTOS NO LIVRO TEXTO DE INVENTÁRIO FLORESTAL DE PÉLLICO NETTO E BRENA (1997, P. 97-300).....	174
QUADRO 31 – EXEMPLO DE DADOS GERADOS PARA O EXERCÍCIO 1 DO “ALUNO 1” .....	176
QUADRO 32 –RESPOSTAS DOS ALUNOS NOS ITENS ABERTOS (11 A 13) DO INSTRUMENTO DE AVALIAÇÃO.....	179
QUADRO 33 – DADOS EM M <sup>3</sup> /HA DE UMA POPULAÇÃO DE <i>PINUS SP</i> .....	180

## LISTA DE SIGLAS

AD	–	Amostragem Dupla
ADDIE	–	<i>Analysis, Design, Development, Implementation and Evaluation</i> (Análise, Projeto, Desenvolvimento, Implantação e Avaliação)
AI	–	Amostragens Independentes
APDIA	–	Análise, Projeto, Desenvolvimento, Implementação e Avaliação
ARP	–	Amostragem com Repetição Parcial
ART	–	Amostragens com Repetição Total
DAP	–	Diâmetro na Altura do Peito
FDD	–	<i>Feature-Driven Development</i> (Desenvolvimento Guiado por Funcionalidades)
FURB	–	Universidade de Blumenau
GRC	–	<i>Generalized Reduced Gradient</i> (Gradiente Reduzido Generalizado)
IDE	–	<i>Integrated Development Enviroment</i> (Ambiente de Desenvolvimento Integrado)
IEEE	–	<i>Institute of Electrical and Electronics Engineers</i> (Instituto de Engenheiros Eletricistas e Eletrônicos)
IFC	–	Inventário Florestal Contínuo
KDE	–	<i>Kernel Density Estimator</i> (Núcleo Estimador de Densidade)
LiDAR	–	<i>Light Detection and Ranging</i> (Detecção e Variação da Luz)
PRNG	–	<i>Pseudo Random Number Generation</i> (Geração de Números Pseudoaleatórios)
SQL	–	<i>Structured Query Language</i> (Linguagem de Consulta Estruturada)
TRNG	–	<i>True Random Number Generation</i> (Geração de Números Verdadeiramente Aleatórios)
UDESC	–	Universidade do Estado de Santa Catarina
UFPR	–	Universidade Federal do Paraná



UFV	–	Universidade Federal de Viçosa
UFV	–	Universidade Federal de Viçosa
UML	–	<i>Unified Modeling Language</i> (Linguagem de Modelagem Unificada)
UnB	–	Universidade de Brasília
UnC	–	Universidade do Contestado
UP	–	<i>Unified Process</i> (Processo Unificado)
URL	–	<i>Uniform Resource Locator</i> (Localizador Uniforme de Recursos)
USP	–	Universidade de São Paulo
UTF	–	Universidade Federal do Tocantins
UTFPR	–	Universidade Tecnológica do Paraná
XP	–	<i>Extreme Programming</i> (Programação Extrema)

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>18</b>
1.1 OBJETIVOS.....	21
1.1.1 Objetivo geral.....	21
1.1.2 Objetivos Específicos .....	22
<b>2 REVISÃO DA LITERATURA.....</b>	<b>23</b>
2.1 PROCESSOS DE AMOSTRAGEM EM INVENTÁRIO FLORESTAL .....	23
2.1.1 Parâmetros da população.....	26
2.1.2 Amostragem aleatória simples.....	32
2.1.3 Amostragem estratificada .....	37
2.1.4 Amostragem sistemática.....	40
2.1.5 Amostragem em dois estágios.....	44
2.1.6 Amostragem em conglomerados .....	47
2.1.7 Amostragem sistemática com múltiplos inícios aleatórios .....	49
2.1.8 Amostragem em múltiplas ocasiões .....	50
2.2 CONSTRUÇÃO DE SOFTWARE E A MODELAGEM ORIENTADA A OBJETOS .....	61
2.2.1 Modelos de processos para construção de <i>software</i> .....	62
2.2.2 Modelagem orientada a objetos.....	69
2.3 GERAÇÃO ALEATÓRIA DE DADOS COM BASE NO MÉTODO NÃO PARAMÉTRICO NÚCLEO ESTIMADOR VARIÁVEL .....	87
2.3.1 Método da inversa da função distribuição acumulada (fda).....	89
2.3.2 Método do núcleo estimador variável .....	93
<b>3 MATERIAL E MÉTODOS.....</b>	<b>98</b>
3.1 METODOLOGIA PARA A CONSTRUÇÃO DO SOFTWARE EDUCACIONAL E BIBLIOTECA DE PROGRAMAÇÃO .....	98
3.1.1 Análise .....	99
3.1.2 Projeto .....	100
3.1.3 Desenvolvimento .....	101
3.1.4 Implantação .....	102
3.1.5 Avaliação .....	102
3.2 TECNOLOGIAS UTILIZADAS .....	104
3.3 MÉTODO PARA GERAÇÃO ALEATÓRIA DE DADOS.....	106
<b>4 RESULTADOS E DISCUSSÃO.....</b>	<b>112</b>
4.1 BIBLIOTECA DE PROGRAMAÇÃO PARA PROCESSOS DE AMOSTRAGEM EM INVENTÁRIO FLORESTAL.....	112
4.1.1 Modelagem da unidade amostral.....	113
4.1.2 Modelagem da unidade amostral de área fixa .....	114
4.1.3 Modelagem da amostra .....	115
4.1.4 Modelagem dos processos de amostragem .....	117
4.1.5 Resultados dos testes da biblioteca de programação .....	130
4.2 APRESENTAÇÃO DO GERADOR DE DADOS PSEUDOALEATÓRIO.....	130
4.3 APRESENTAÇÃO DO SISTEMA .....	134
4.3.1 Desafios do grupo 1.....	139
4.3.2 Desafios do grupo 2.....	141
4.3.3 Desafios do grupo 3.....	144

4.3.4	Desafios do grupo 4.....	146
4.3.5	Sistema de Pontuação.....	147
4.4	AVALIAÇÃO DO SOFTWARE.....	148
5	CONCLUSÕES.....	152
6	RECOMENDAÇÕES PARA TRABALHOS FUTUROS.....	154
	REFERÊNCIAS.....	155
APÊNDICE 1	– ESTUDO DOS MOMENTOS DA FUNÇÃO NÚCLEO ESTIMADOR VARIÁVEL.....	162
APÊNDICE 2	– EXEMPLO DE USO DO MÉTODO NÚCLEO ESTIMADOR VARIÁVEL.....	165
APÊNDICE 3	– DIAGRAMA ENTIDADE RELACIONAMENTO.....	170
APÊNDICE 4	– QUADRO DAS CLASSES DA BIBLIOTECA DE PROGRAMAÇÃO PARA OS PROCESSOS DE AMOSTRAGEM.....	171
APÊNDICE 5	– EVIDÊNCIA DOS TESTES REALIZADOS NA BIBLIOTECA DE PROGRAMAÇÃO PARA OS PROCESSOS DE AMOSTRAGEM.....	173
APÊNDICE 6	– EXEMPLO DE DADOS GERADOS PARA O EXERCÍCIO 1 DO “ALUNO 1”.....	176
APÊNDICE 7	– TESTE DE KOLMOGOROV-SMIRNOV PARA 102 CONJUNTOS DE DADOS GERADOS.....	178
APÊNDICE 8	– RESPOSTA DOS ALUNOS AOS ITENS ABERTOS DO INTRUMENTO DE AVALIAÇÃO.....	179
ANEXO 1	– QUADRO COM DADOS ORIGINAIS DO LIVRO TEXTO DE PÉLLICO NETTO E BRENA (1997) .....	180

## 1 INTRODUÇÃO

O Inventário Florestal é uma importante disciplina para a Engenharia Florestal, e também para subsidiar a definição de políticas de desenvolvimento, uso e conservação dos recursos florestais. Por meio do inventário florestal é possível conhecer a quantidade e a qualidade dos recursos florestais em uma determinada área.

Os conceitos estatísticos envolvidos em um inventário florestal, mais especificamente nos conteúdos dos processos de amostragem, são muito importantes para um aluno de Engenharia Florestal. A Resolução nº 3, de 2 de fevereiro de 2006, do Conselho Nacional de Educação (BRASIL, 2006), estabelece as Diretrizes Curriculares Nacionais para o curso de graduação em Engenharia Florestal e determina que o conteúdo da disciplina de Inventário Florestal deve fazer parte do núcleo de conteúdos profissionais essenciais. Todos os cursos de Engenharia Florestal, pesquisados em nove universidades, por meio da análise de suas ementas na internet, revelaram que possuem os métodos e processos de amostragem em suas ementas na disciplina de Inventário Florestal. As universidades pesquisadas foram: FURB (Universidade de Blumenau), UDESC (Universidade do Estado de Santa Catarina), UFPR (Universidade Federal do Paraná), UTF (Universidade Federal do Tocantins), UFV (Universidade Federal de Viçosa), UnB (Universidade de Brasília), UnC (Universidade do Contestado), USP (Universidade de São Paulo) e UTFPR (Universidade Tecnológica Federal do Paraná).

A aplicação de tecnologia da informação no processamento e na coleta de dados dos inventários florestais é uma realidade inegável, tanto na academia, como nas empresas. *Softwares* para o processamento do inventário florestal são aplicados para otimizar os custos e o tempo de um inventário florestal. Alguns exemplos de *softwares* que podem ser utilizados para auxiliar no processamento do inventário florestal são: Mata Nativa (CIENTEC, 2014), Faça Floresta (INSTITUTO BRASILEIRO

DE FLORESTAS, 2014), SifcubCE, SifcubDesktop, Sifcub, Sifprog, Sifplan, SifcadSifquality (TREESOFTWARE, 2013) e C7 LDFP (PROJETO CR CAMPEIRO - UFSM, 2014).

As tecnologias voltadas para o ensino, contudo, não evoluíram na mesma velocidade. No ensino superior e especialmente nas engenharias, os conteúdos ministrados possuem uma alta complexidade, com uma teoria bastante extensa. Tais conteúdos normalmente são ministrados por meio de aulas expositivas e resolução de exercícios. Esse quadro pode ficar ainda pior quando se leva em conta que a geração atual de estudantes cresceu em meio ao mundo digital: computadores, internet, videogames, telefones celulares, *tablets* e outros instrumentos eletrônicos. Simuladores computacionais, conteúdos multimídia e jogos digitais podem ser desenvolvidos para apresentar tópicos de engenharia de modo a superar algumas limitações do formato expositivo (DESHPANDE e HUANG, 2009).

Pensando nessas ferramentas, os objetos de aprendizagem se tornaram um conceito bastante difundido e aceito pelos educadores nas últimas décadas. Segundo Wiley (2000), “um objeto de aprendizagem pode ser definido como qualquer recurso digital que possa ser reutilizado para o suporte ao ensino”. Kay e Knaack (2007) colocaram mais especificamente que um objeto de aprendizagem pode ser definido como “uma ferramenta interativa baseada na web que suporta o aprendizado de conceitos específicos por meio do melhoramento, amplificação e orientação do processo cognitivo dos estudantes”.

O Ministério da Educação do Brasil disponibiliza um portal na internet para busca e catalogação de objetos de aprendizagem nacionais e internacionais, que estejam disponíveis de forma gratuita. Uma busca nesse portal revelou que existem 15 objetos de aprendizagem na área de Recursos Florestais e Engenharia Florestal. Foram encontrados 4 vídeos, 9 imagens e 1 animação. Nenhum *software* educacional foi encontrado (BRASIL, 2008).

Existem muitos estudos a respeito da aplicação de *softwares* educacionais no ensino superior de diversas áreas, tais como: engenharia civil (BARRETTO,

PIAZZALUNGA e RIBEIRO, 2003), ciência da computação (NUGENT, SOH e SAMAL, 2006), engenharia eletrônica (ROSELLÓ *et al.*, 2014), engenharia da computação (DJORDJEVIC, NIKOLIC e MILENKOVIC, 2005) e engenharia mecânica (WANG *et al.*, 2006) (MUÑOZ-ABELLA, ÁLVAREZ-CALDAS e RUBIO, 2011). Não foi encontrada na literatura pesquisada qualquer relato sobre construção e aplicação de *software* educacional na área de engenharia florestal.

Deshpande e Huang (2009), em seu artigo de revisão, cria uma taxonomia para jogos de simulação aplicados à cursos de engenharia. Eles utilizam essa taxonomia para fazer uma pesquisa sobre *softwares* educacionais de simulação. As grandes áreas da engenharia citadas nessa taxonomia são: engenharia civil, engenharia elétrica, engenharia da computação, engenharia química, engenharia mecânica, engenharia industrial e engenharia ambiental. Para cada uma das áreas dessa taxonomia, esses autores citaram alguns trabalhos que descrevem *softwares* educacionais de simulação. Na área de engenharia ambiental (área correlata a engenharia florestal) eles citaram alguns trabalhos desenvolvidos com os seguintes temas: lixo industrial (HIROSE, SUGIURA e SHIMOMOTO, 2004), gerenciamento da qualidade da água (JEHNG-JUNG e YING-JIE, 1996), sustentabilidade (TORRES e MACEDO, 2000) e efeitos ecológicos e econômicos do consumo de alimentos (HANSMANN *et al.*, 2005).

Segundo Deshpande e Huang (2009) uma grande variedade de *softwares* educacionais de simulação tem sido desenvolvida, por outro lado observa-se na revisão realizada por eles, que não foi encontrado *softwares* especificamente na área de Engenharia Florestal.

Este trabalho apresenta um *software* educacional como um objeto de aprendizagem, que foi desenvolvido no programa de pós-graduação de Engenharia Florestal da UFPR, com o objetivo de ajudar alunos dos cursos de graduação e pós-graduação na compreensão do conteúdo dos processos de amostragem da disciplina de Inventário Florestal.

O *software* desenvolvido é executado em um navegador, acessado pela internet; ele sugere uma série de exercícios relacionados aos processos de amostragem; nele o objetivo do aluno é calcular os parâmetros das populações e as estimativas das amostras. O *software* cria problemas personalizados para cada aluno, permite o uso de uma planilha eletrônica para a realização dos cálculos, corrige os valores inseridos pelo aluno e fornece uma pontuação por acerto. O principal objetivo foi criar um ambiente no qual o aluno possa exercitar os conceitos vistos em sala de aula sem depender do professor para realizar as correções.

Para a construção desse *software* educacional foram estudados os processos de amostragem florestal e, como consequência, uma biblioteca de programação foi modelada e implementada utilizando os conceitos de orientação a objetos.

Além da biblioteca de programação, foi utilizado um método não paramétrico, chamado “núcleo estimador variável”, para gerar dados pseudoaleatórios baseados em uma distribuição de probabilidade desconhecida.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo geral

Desenvolver e avaliar um *software* educacional para a disciplina de Inventário Florestal, como um recurso didático acessado pela Internet, no qual os alunos de graduação e pós-graduação possam exercitar os processos de amostragem, que são um conteúdo fundamental para o curso.

### 1.1.2 Objetivos Específicos

- a) Modelar, implementar e testar uma biblioteca de programação orientada a objetos, na linguagem Java, que realize os cálculos estatísticos de processos de amostragem;
- b) Construir e testar um gerador de dados capaz de produzir dados aleatórios, que respeitem uma determinada distribuição, para que cada aluno tenha seu próprio conjunto de dados nos exercícios propostos pelo sistema;
- c) Avaliar o *software* gerado com uma turma de graduação em Engenharia Florestal da Universidade Federal do Paraná.



## 2 REVISÃO DA LITERATURA

Este capítulo descreve os principais conceitos utilizados na construção e avaliação do *software* educacional para os processos de amostragem florestal.

A seção 2.1 faz uma revisão dos processos de amostragem que foram implementados no *software*, mais especificamente na biblioteca de programação apresentada na seção 4.1.

A seção 2.2 apresenta os principais conceitos dos processos e métodos de modelagem utilizados na construção do *software* educacional e da biblioteca de programação para os processos de amostragem florestal. Na construção do *software* foram utilizadas algumas práticas dos modelos de processos de construção de *software*, fazendo-se necessária uma revisão dos principais modelos (seção 2.2.1). Na modelagem e construção da biblioteca de programação (seção 4.1) foi utilizado o paradigma orientado a objetos. Por esse motivo, a seção 2.2.2 apresenta uma revisão bibliográfica do tema. A revisão bibliográfica dos temas abordados na seção 2.2 não tem a intenção de apresentar todos os detalhes dos processos e métodos de modelagem utilizados para a construção do *software*, mas os conceitos principais, a fim de que o leitor tenha informações para os capítulos e seções subsequentes.

A geração aleatória de dados, baseada no método não paramétrico “núcleo estimador variável”, foi utilizada para possibilitar a criação de um conjunto individual de dados para os usuários do sistema. A seção 2.3, por sua vez, descreve esse método estatístico e suas utilizações na área florestal.

### 2.1 PROCESSOS DE AMOSTRAGEM EM INVENTÁRIO FLORESTAL

Segundo Péllico Netto e Brena (1997, p. 6), “inventário florestal é uma atividade que visa obter informações qualitativas e quantitativas dos recursos

florestais existentes em uma área pré-especificada”. Por outro lado, segundo Sanquetta *et al.* (2009, p. 2), esse conceito é genérico o bastante para incluir qualquer tipo de levantamento florestal, mas salienta que a “marca registrada” de um inventário é a sua representatividade amostral e sua validade estatística. Laar e Akça (2007, p. 229) acrescentam que o objetivo de um inventário florestal é obter informações quantitativas e qualitativas sobre os recursos da floresta e seu ambiente físico, em um ponto específico no tempo, por um custo razoável. Para Queiroz (2012, p. 19), “inventário florestal é o ramo da ciência florestal que visa avaliar as variáveis qualitativas e quantitativas da floresta e suas inter-relações, assim como dinâmicas de crescimento e sucessão florestal, servindo de base para a formulação de planos de utilização de produtos florestais, manejo sustentado integrado da floresta, bem como para alicerçar propostas de planos de desenvolvimento e política florestal de caráter regional ou nacional.”

Péllico Netto e Brena (1997, p. 6-11) classificam um inventário florestal segundo seus objetivos, abrangência, forma de obtenção de dados, abordagem da população no tempo e grau de detalhamento dos seus recursos. Considerando a classificação de acordo com a forma de obtenção de dados, esses autores afirmam que se podem agrupar os inventários florestais nos seguintes tipos: inventário por enumeração total ou censo, inventário por amostragem e inventários por tabela de produção.

Sanquetta *et al.* (2009, p. 4) afirmam que o inventário por enumeração total ou por censo pressupõe a medida de todos os indivíduos da população. Porém, devido ao alto custo, esse tipo é restrito às aplicações específicas, como no caso de florestas cujo valor econômico, científico ou cultural seja muito elevado, ou por exigências legais, ou ainda no caso de uma área de pequenas dimensões.

A maior parte dos inventários florestais está no grupo dos efetuados por amostragem, nos quais é necessário medir apenas alguns indivíduos da população para formar uma amostra. As técnicas de amostragem são comumente utilizadas para estimar as informações quantitativas e qualitativas da floresta. Essas técnicas

otimizam o tempo e o custo dos inventários florestais se comparadas ao censo ou enumeração completa da população (HUSH, MILLER e BEERS, 1982, p. 150-157) (PÉLLICO NETTO e BRENA, 1997, p. 14-15) (LAAR e AKÇA, 2007, p. 230-232) (SANQUETTA *et al.*, 2009, p. 4).

Segundo Druszcz *et al.* (2013), as técnicas de amostragem em campo para coleta de informações quantitativas e qualitativas da floresta dificilmente serão abolidas, mesmo levando-se em consideração as inovações tecnológicas, como, por exemplo, a tecnologia *LiDAR (Light Detection and Ranging)*. Para esses autores, contudo, também existe um custo na aplicação dessas técnicas, o qual aumenta conforme o nível de precisão e a velocidade requerida.

Três classificações para as técnicas de amostragem são apresentadas por Péllico Netto e Brena (1997, p. 24-28): segundo a periodicidade, a estrutura e a abordagem da população. Nessa última, classificam as técnicas de amostragem como métodos de amostragem e processos de amostragem.

Os métodos de amostragem, segundo Péllico Netto e Brena (1997, p. 26), se referem à seleção dos indivíduos que participam da unidade amostral, como por exemplo, método da área fixa, método de Bitterlich, método de Prodan, método de Strand e método 3-P. Sanquetta *et al.* (2009, p. 47) complementam que os métodos de amostragem podem ser considerados mediante a utilização de métodos de área fixa e métodos de área variável.

Por outro lado, os processos de amostragem se referem à abordagem do conjunto das unidades amostrais, ou seja, de uma amostra da população formada por unidades amostrais. Os processos de amostragem, para uma única ocasião, podem ser classificados em: aleatório, sistemático ou misto. A categoria dos processos para múltiplas ocasiões pode ser classificada em: amostragens independentes, amostragem com repetição total, amostragem dupla e amostragem com repetição parcial (PÉLLICO NETTO e BRENA, 1997, p. 26-28).

Nas próximas seções serão discutidos os processos de amostragem que foram implementados na biblioteca de programação construída (ver seção 4.1) e que estão dispostos como exercícios no *software* educacional (ver seção 4.1).

### 2.1.1 Parâmetros da população

A enumeração total ou censo, apesar de ser classificada como um tipo de inventário florestal restrito a aplicações específicas (SANQUETTA *et al.*, 2009, p. 4), para fins didáticos pode ser utilizada para comparar os parâmetros com as estimativas obtidas pelos diversos processos de amostragem. Por esse motivo, esses processos com parâmetros da população foram implementados no *software* educacional apresentado na seção 4.1 e serão revisados nessa seção.

A população é definida por Hush *et al.* (1982, p. 156) como a agregação de todas as possíveis unidades amostrais e complementa que as unidades amostrais podem ser parcelas, unidades administrativas, áreas fixas, faixas ou pontos. Para Köhl *et al.* (2006, p. 71), uma população compreende todos os elementos onde a amostra é obtida e complementa que a população é também unicamente definida em termos de unidades amostrais como a união de todas as possíveis amostras dessas unidades. Péllico Netto e Brena (1997, p. 15) definem “população” como um conjunto de seres da mesma natureza, que ocupam um determinado espaço em um determinado tempo. Cochran (1977, p. 5) afirma que a palavra “população” é usada para denotar a agregação de onde a amostra é escolhida. Esse autor complementa que a população que será amostrada (população amostrada) pode coincidir com a população da qual se deseja a informação (população objetivo), mas que, em alguns casos, por razões de conveniência ou praticidade, a população amostrada é mais restrita do que a população objetivo.

Laar e Akça (2007, p. 253) afirmam que as amostragens baseadas em árvores individuais como unidades amostrais podem ser utilizadas para estimar a média da

densidade de madeira ou a idade média dentro de uma parcela. Mas afirmam também que não é recomendado utilizar árvores individuais como unidades amostrais em inventários florestais para estimar área basal, número de árvores e volume por hectare.

As próximas seções apresentam os quatro tipos de processos de amostragem baseados nos parâmetros da população.

### Parâmetros da população por unidade amostral

Os parâmetros da população podem ser medidos quando todas as unidades amostrais da população forem medidas. Apresentam-se, no Quadro 1, os parâmetros da população, segundo Péllico Netto e Brena (1997, p. 99-100).

Parâmetro	Fórmula
Média	$\bar{X} = \frac{\sum_{i=1}^N X_i}{N}$
Variância	$S_x^2 = \frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}$
Desvio padrão	$S_x = \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{N - 1}}$
Coeficiente de variação	$CV = \frac{S_x}{\bar{X}} \cdot 100$
Total da população	$X = N \cdot \bar{X}$
<b>Notação</b> $\bar{X}$ = Média da população $N$ = Quantidade total de unidades amostrais da população $X_i$ = Variável de interesse da unidade amostral $i$ $S_x^2$ = Variância da população $S_x$ = Desvio padrão da população $CV$ = Coeficiente de variação $X$ = Total da população	

QUADRO 1 – PARÂMETROS DA POPULAÇÃO POR UNIDADE AMOSTRAL  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 99,100)

### Parâmetros da população estratificada

A população pode ser estratificada por algum critério, como, por exemplo, pela idade em que o povoamento foi plantado. Para cada estrato ( $h$ ) podem-se calcular os parâmetros apresentados no Quadro 1.

Segundo Péllico Netto e Brena (1997, p. 100-101), os parâmetros para a população estratificada foram formulados e estão apresentados no Quadro 2.

Parâmetro	Fórmula
Média estratificada	$\bar{X}_{st} = \frac{\sum_{h=1}^L N_h \cdot \bar{X}_h}{N}$
Variância estratificada	$S_{st}^2 = \frac{\sum_{h=1}^L N_h \cdot S_h^2}{N}$
Variância entre os estratos	$S_e^2 = \frac{\sum_{h=1}^L N_h \cdot (\bar{X}_h - \bar{X})^2}{N - 1}$
Variância dentro dos estratos	$S_d^2 = \frac{\sum_{h=1}^L \sum_{i=1}^{N_h} (X_{ih} - \bar{X}_h)^2}{N - L}$
Total da população por estrato	$X_h = N_h \cdot \bar{X}_h$
<b>Notação</b> $\bar{X}_{st}$ = Média estratificada $N_h$ = Quantidade total de unidades amostrais do estrato $h$ $N$ = Quantidade total de unidades amostrais da população $\bar{X}_h$ = Média do estrato $h$ $S_{st}^2$ = Variância estratificada $S_h^2$ = Variância do estrato $h$ $S_e^2$ = Variância entre estratos $\bar{X}$ = Média da população $S_d^2$ = Variância dentro dos estratos $L$ = Quantidade total de estratos $X_{ih}$ = Variável de interesse da unidade amostral $i$ do estrato $h$ $X_h$ = Total da população do estrato $h$	

QUADRO 2 – PARÂMETROS DA POPULAÇÃO ESTRATIFICADA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 100,101)

### Parâmetros da população por faixa

Nesse caso, a população é estratificada por faixas, considerando-se unidades contíguas da população. Para cada faixa ( $j$ ), podem-se calcular os parâmetros apresentados no Quadro 1.

Péllico Netto e Brena (1997, p. 101-102) definem a formulação dos parâmetros da população por faixa que foi adaptada e apresentada no Quadro 3.

Parâmetro	Fórmula
Média estratificada das faixas	$\bar{X}_{fx} = \frac{\sum_{f=1}^F N_f \cdot \bar{X}_f}{F}$
Variância estratificada das faixas	$S_{fx}^2 = \frac{\sum_{f=1}^F (\bar{X}_f - \bar{X}_{fx})^2}{F-1}$
Desvio padrão estratificado das faixas	$S_{fx} = \sqrt{\frac{\sum_{f=1}^F (\bar{X}_f - \bar{X}_{fx})^2}{F-1}}$
Coeficiente de variação	$CV = \frac{S_{fx}}{\bar{X}_{fx}} \cdot 100$
<b>Notação</b> $\bar{X}_{fx}$ = Média estratificada das faixas $F$ = Quantidade total de faixas $N_f$ = Quantidade total de unidades amostrais por faixa $f$ $\bar{X}_f$ = Média da faixa $f$ $S_{fx}^2$ = Variância estratificada das faixas $S_{fx}$ = Desvio padrão estratificado das faixas $CV$ = Coeficiente de variação	

QUADRO 3 – PARÂMETROS DA POPULAÇÃO POR FAIXA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 101,102)

### Parâmetros da população por faixas com tamanhos diferentes

A população pode ser estratificada por faixas, mas com a particularidade das faixas possuírem tamanhos diferentes, ou seja, número de subunidades amostrais diferentes. Para cada faixa ( $f$ ), podem-se calcular os parâmetros apresentados no Quadro 1.

Segundo Péllico Netto e Brena (1997, p. 104), a aplicação de estimativas de razão é o método estatístico utilizado para análise de amostragem com unidades de tamanhos diferentes. Esses autores definem a formulação estatística para esse método, conforme apresentada no Quadro 4.

Parâmetro	Fórmula
Razão média	$\bar{R} = \frac{\sum_{f=1}^F X_f}{\sum_{f=1}^F Y_f}$

Parâmetro	Fórmula
Variância da razão	$S_R^2 = \frac{\bar{R}^2}{F-1} \left[ \frac{\sum_{f=1}^F X_f^2}{\bar{X}_{fx}^2} + \frac{\sum_{f=1}^F Y_f^2}{\bar{Y}_{fx}^2} - \frac{2 \sum_{f=1}^F X_f \cdot Y_f}{\bar{X}_{fx} \cdot \bar{Y}_{fx}} \right]$
Desvio padrão da razão	$S_R = \sqrt{S_R^2}$
Coeficiente de variação	$CV = \frac{S_R}{\bar{X}_{fx}} \cdot 100$
<b>Notação</b> $\bar{R}$ = Razão média da população $F$ = Quantidade total de faixas $Y_f$ = Tamanho em área da faixa $f$ $X_f$ = Volume da faixa $f$ $S_R^2$ = Variância da razão da população $\bar{X}_{fx}$ = Média dos volumes das faixas (conforme Quadro 3) $\bar{Y}_{fx}$ = Média das áreas das faixas $S_R$ = Desvio padrão da razão da população $CV$ = Coeficiente de variação	

QUADRO 4 – PARÂMETROS DA POPULAÇÃO POR FAIXA COM TAMANHOS DIFERENTES  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 105)

### Parâmetros da população por unidade primária

Considera-se a subdivisão da população em unidades primárias, com um número fixo de unidades secundárias cada uma. O Quadro 5 apresenta um exemplo fictício de uma população de *Pinus* sp com 3,6 ha de área, constituída de 36 unidades secundárias de 0,1 ha, cujos volumes estão expressos por hectare.

	a	b	c	d	e	f	
1	80	92	96	94	90	85	A
2	99	69	102	103	91	123	
3	86	69	85	127	98	102	
4	81	89	122	110	80	99	B
5	131	115	92	76	136	157	
6	162	100	90	76	116	83	
	I			II			

QUADRO 5 – EXEMPLO DE UMA POPULAÇÃO COM UNIDADES PRIMÁRIAS E SECUNDÁRIAS  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 98)

No Quadro 5 tem-se quatro unidades primárias, representadas pela combinação das letras em maiúsculo (“A” e “B”) e dos números romanos (I e II), sendo elas: “AI”, “AII”, “BI” e “BII”. As unidades secundárias estão representadas pela combinação dos números arábicos (de 1 até 6) e das letras minúsculas (de “a” até “f”), totalizando 36 unidades secundárias. Cada unidades primária possui seis unidades



secundárias. Por exemplo, as unidades secundárias da unidade primária “A1” são: 1a, 1b, 1c, 2a, 2b, 2c, 3a, 3b e 3c. Os respectivos valores em volume/hectare dessas unidades secundárias são: 80, 92, 96, 99, 69, 102, 86, 69 e 85.

O Quadro 6 utiliza a representação “i” para numeração das unidades amostrais primárias e “j” para representação das unidades amostrais secundárias dentro das unidades primárias. Onde “i” e “j” são números inteiros. Para o exemplo do Quadro 5, pode-se afirmar que a unidade amostral primária “A1”, pode ser mapeada para “i=1”, a unidade amostral “A11” pode ser mapeada para “i=2” e assim por diante. Quando queremos especificar uma unidade amostral secundária define-se primeiro a unidade amostral primária ao qual a unidade secundária pertence e depois um número inteiro “i” para representar a unidade secundária dentro da primária. Por exemplo, se quisermos representar no Quadro 6 a unidade secundária 5e, devemos identificar a sua unidade primária, que nesse caso será B11, ou i=4. Em seguida identificamos a ordem da unidade secundária dentro da unidade primária, nesse caso “j=5”.

Péllico Netto e Brena (1997, p. 104) definem a formulação dos parâmetros da variância entre os estratos e dentro dos estratos de uma população por unidade primária conforme apresentada no Quadro 6.

Parâmetro	Fórmula
Quadrado médio dentro	$MQ_{dentro} = \frac{\sum_{i=1}^P \sum_{j=1}^S (X_{ij} - \bar{X}_i)^2}{P \cdot (S - 1)}$
Variância dentro	$S_d^2 = MQ_{dentro}$
Quadrado médio entre unidade primárias	$MQ_{entre} = \frac{\sum_{i=1}^P S \cdot (\bar{X}_i - \bar{X})^2}{P - 1}$
Variância entre unidades primárias	$S_e^2 = \frac{MQ_{entre} - MQ_{dentro}}{S}$
<b>Notação</b> $MQ_{dentro}$ = Quadrado Médio Dentro $P$ = Quantidade de unidades primárias $S$ = Quantidade de subunidades em cada unidade primária $X_{ij}$ = Variável de interesse da subunidade amostral $j$ da unidade amostral primária $i$ $S_d^2$ = Variância dentro $MQ_{entre}$ = Quadrado médio entre unidades primárias $\bar{X}_i$ = Média das subunidades da unidade primária $i$ $\bar{X}$ = Média das unidades primárias	

**Notação** $S_p^2$  = Variância entre unidades primárias

QUADRO 6 – PARÂMETROS DA POPULAÇÃO POR UNIDADE PRIMÁRIA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 106)

### 2.1.2 Amostragem aleatória simples

Vários autores concordam que o processo de amostragem aleatória simples é base para praticamente qualquer outro processo amostral (HUSH, MILLER e BEERS, 1982, p. 162) (SANQUETTA *et al.*, 2009, p. 84) (PÉLLICO NETTO e BRENA, 1997, p. 107) (PRODAN *et al.*, 1997, p. 207). Esses autores afirmam também que, nesse tipo de amostragem, todas as combinações possíveis de (n) unidades amostrais da população devem ter igual chance de participar da amostra e que a seleção de cada unidade amostral deve ser livre de qualquer escolha e totalmente independente da seleção das demais unidades amostrais. Köhl *et al.* (2006, p. 71) complementam que o princípio geral da amostragem é selecionar um subconjunto de unidades (entenda-se, uma amostra) de uma população, para mensurar intensamente esse subconjunto e desenhar inferências da amostra para a população toda. Asseveram que esse processo é apropriado quando qualquer unidade da população tem igual chance de ser selecionada. Cochran (1977, p. 18) define a amostragem aleatória simples como sendo um método para seleção de  $n$  unidades de um conjunto de  $N$ , tal que cada uma das  $C_n^N$  amostras distintas tenha igual chance de ser desenhada.

A definição de amostra também é apresentada por diversos autores. Hush *et al.* (1982, p. 156) afirmam que o grupo de unidades amostrais escolhidos para serem medidas constituem a amostra. Para Köhl *et al.* (2006, p. 72) uma amostra consiste de um número de unidades amostrais (ou simplesmente “unidades”) selecionadas de uma população por algum processo. Péllico Netto e Brena (1997, p. 16) afirmam que uma amostra pode ser definida como uma parte da população, constituída de

indivíduos que apresentam características comuns que identificam a população a que pertencem.

A amostragem aleatória simples pode ser realizada com e sem repetição. Nas amostragens com repetição cada unidade amostral pode ser selecionada mais de uma vez. Péllico Netto e Brena (1997, p. 108) afirmam que a maioria das amostras, utilizando parcelas de área fixa ou faixas, em inventários florestais, é selecionada sem reposição. Por outro lado, afirmam que, se forem usados pontos amostrais, a população é infinita e a amostra obtida é equivalente à seleção com reposição. Nas amostragens sem repetição as unidades amostrais que fazem parte da amostra são selecionadas uma única vez. Nesse tipo, para populações finitas, pode-se definir a quantidade de possíveis amostras que podem ser formadas selecionando  $n$  unidades amostrais de uma população com um total de  $N$  unidades, conforme apresentado na equação 1 (KÖHL, MAGNUSSEN e MARCHETTI, 2006, p. 75) (PÉLLICO NETTO e BRENA, 1997, p. 107) (COCHRAN, 1977, p. 18).

$$C_n^N = \frac{N!}{n!(N-n)!} \quad (1)$$

Cochran (1977, p. 18) acrescenta que a unidade na população é numerada de 1 até  $N$  e uma série de números randômicos entre 1 e  $N$  é gerada por um computador, por exemplo. A geração de números randômicos será revisada na seção 2.3 deste trabalho.

Segundo Sanquetta *et al.* (2009, p. 95) se forem utilizadas unidades amostrais de área fixa, ou seja, se a população for finita, o desenho da amostra na área florestal a ser inventariada pode ser combinada de várias maneiras. Se uma população tiver um número total de ( $N$ ) unidades, o número de combinações possíveis de ( $n$ ) unidades a serem amostradas na população é dado pela equação 1.

O Quadro 7 apresenta as principais estimativas obtidas por meio da amostragem aleatória simples, definidas por Péllico Netto e Brena (1997, p. 109-114).

Estimativa	Fórmula
Média aritmética	$\bar{x} = \frac{\sum_{i=1}^n X_i}{n}$
Variância	$s_x^2 = \frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n-1}$
Desvio Padrão	$s_x = \sqrt{\frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n-1}}$
Variância da média	$s_{\bar{x}}^2 = \frac{s_x^2}{n} \cdot \left( \frac{N-n}{N} \right)$
Fração de amostragem	$f = \frac{n}{N}$
Erro padrão	$s_{\bar{x}} = \pm \frac{s_x}{\sqrt{n}} \cdot \sqrt{(1-f)}$
Coeficiente de variação	$cv = \frac{s_x}{\bar{x}} \cdot 100$
Variância da média relativa	$v_{\bar{x}}^2 = \frac{(cv \cdot \bar{x})^2}{n} \cdot (1-f)$
Erro de amostragem absoluto	$E_a = \pm t \cdot s_{\bar{x}}$
Erro de amostragem relativo	$E_r = \pm \frac{t \cdot s_{\bar{x}}}{\bar{x}}$
Intervalo de confiança para a média	$IC[\bar{x} - t \cdot s_{\bar{x}} \leq \bar{X} \leq \bar{x} + t \cdot s_{\bar{x}}] = P$
Total da população	$\hat{X} = N \cdot \bar{x}$
Intervalo de confiança para o total	$IC[\hat{X} - N \cdot t \cdot s_{\bar{x}} \leq X \leq \hat{X} + N \cdot t \cdot s_{\bar{x}}] = P$
Razão média	$\bar{r} = \frac{\bar{x}}{\bar{y}} = \frac{\sum_{i=1}^n X_i}{\sum_{i=1}^n Y_i}$
Variância da razão	$s_r^2 = \frac{\bar{r}^2}{n-1} \cdot \left[ \frac{\sum_{i=1}^n X_i^2}{\bar{x}^2} + \frac{\sum_{i=1}^n Y_i^2}{\bar{y}^2} - \frac{2 \sum_{i=1}^n X_i \cdot Y_i}{\bar{x} \cdot \bar{y}} \right]$
<b>Notação</b> $\bar{x}$ = Estimativa da média $n$ = Número de unidades amostradas $X_i$ = Variável de interesse da unidade amostral $i$ $s_x^2$ = Estimativa da variância $s_x$ = Estimativa do desvio padrão $s_{\bar{x}}^2$ = Estimativa da variância da média $v_{\bar{x}}^2$ = Estimativa da variância da média relativa $N$ = Quantidade total de unidades amostrais da população $f$ = Fração de amostragem $s_{\bar{x}}$ = Estimativa do erro padrão $cv$ = Estimativa do coeficiente de variação $E_a$ = Erro de amostragem absoluto $t$ = Valor bicaudal da distribuição de Student para $n_e$ graus de liberdade e probabilidade de $1-P$	

**Notação** $P$  = Probabilidade de confiança $E_r$  = Erro de amostragem relativo $\hat{X}$  = Estimativa do total da população $\bar{r}$  = Estimativa da razão média $\bar{y}$  = Estimativa da média das áreas das unidades amostrais $X_i$  = Volume da unidade amostral  $i$  $Y_i$  = Área da unidade amostral  $i$  $s_r^2$  = Estimativa da variância da razão

QUADRO 7 – ESTIMATIVAS PARA AMOSTRAGEM ALEATÓRIA SIMPLES  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 109-114)

A determinação do número de unidades amostrais que fará parte de uma amostra é muito importante em qualquer processo de amostragem. A relação entre a quantidade de unidades amostrais de uma amostra e a quantidade total de unidades amostrais da população pode ser definida como “intensidade de amostragem”, que está apresentada no Quadro 7 como fração de amostragem.

Segundo Cochran (1977, p. 72) amostras muito grandes implicam em um desperdício de recursos, e amostras muito pequenas podem comprometer a utilidade dos resultados. Segundo Laar e Akça (2007, p. 256) a estimativa do número de unidades amostrais deve garantir que um dado erro máximo não será excedido com uma probabilidade específica igual ou menor do que  $(1 - P)$ .

Péllico Netto e Brena (1997, p. 115-117) afirmam que a intensidade de amostragem pode ser calculada em função da variância ou do coeficiente de variação, e apresentam as fórmulas específicas para o caso de população finita ou infinita. Esses autores afirmam que a diferenciação estatística de população finita e infinita é feita pelo valor do fator de correção  $(1-f)$ , em que  $f$  é a intensidade de amostragem dada pela fórmula do Quadro 7. O Quadro 8 apresenta as fórmulas específicas de intensidade de amostragem para os quatro casos.

	<b>Intensidade de amostragem em função da variância</b>	<b>Intensidade de amostragem em função do coeficiente de variação</b>
<b>População finita</b>	$n = \frac{1}{\frac{(LE \cdot \bar{x})^2}{t^2 \cdot s_x^2} + \frac{1}{N}}$	$n = \frac{t^2 \cdot (cv\%)^2}{(LE\%)^2 + \frac{t^2 \cdot (cv\%)^2}{N}}$
<b>População infinita</b>	$n = \frac{t^2 \cdot s_x^2}{(LE \cdot \bar{x})^2}$	$n = \frac{t^2 \cdot (cv\%)^2}{(LE\%)^2}$

**Notação** $n$  = Intensidade de amostragem $LE$  = Limite do erro de amostragem admitido em percentual $\bar{x}$  = Estimativa da média $t$  = Valor bicaudal da distribuição de *Student* para  $n_a - 1$  graus de liberdade e probabilidade de  $1 - P$  $P$  = Probabilidade de confiança $s_x^2$  = Estimativa da variância $cv\%$  = Estimativa do coeficiente de variação em percentual $N$  = Quantidade total de unidades

QUADRO 8 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM ALEATÓRIA SIMPLES  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 115-117)

Segundo Péllico Netto e Brena (1997, p. 116, 117) depois de obtida a intensidade de amostragem, é necessário proceder a um ajuste, pois o número de unidades amostrais que foi utilizado para obtê-la é arbitrário e proveniente de um inventário piloto. A Figura 1 demonstra o algoritmo aplicado no *software* desenvolvido neste trabalho para obter a intensidade de amostragem em função da variância. Para se calcular em função do coeficiente de variação, basta trocar as respectivas fórmulas conforme descritas no Quadro 8. As variáveis utilizadas no algoritmo apresentado na Figura 1 estão definidas no Quadro 8, com exceção da variável  $n_a$  que representa a intensidade de amostragem anterior.

**faça** $n_a \leftarrow n$  ;**se**  $(1-f) < 0,98$  **então**

$$n = \frac{1}{\frac{(LE \cdot \bar{x})^2}{t^2 \cdot s_x^2} + \frac{1}{N}} ;$$

**fim-se****se**  $(1-f) \geq 0,98$  **então**

$$n = \frac{t^2 \cdot s_x^2}{(LE \cdot \bar{x})^2} ;$$

**fim-se****enquanto**  $n_a \neq n$ **retorne**  $n$ ;

FIGURA 1 – ALGORITMO PARA CÁLCULO DA INTENSIDADE DE AMOSTRAGEM  
 FONTE: O AUTOR (2014)

### 2.1.3 Amostragem estratificada

De acordo com Cochran (1977, p. 89), na amostragem estratificada a população de  $N$  unidades amostrais é primeiramente dividida em subpopulações de  $N_1, N_2, \dots, N_L$  unidades respectivamente. As subpopulações não são sobrepostas e juntas compreendem o total da população, portanto  $N_1 + N_2 + \dots + N_L = N$ . Ainda, segundo esse autor, as subpopulações são chamadas de estratos.

Köhl *et al.* (2006, p. 105) estabelecem que a estratificação permite que se forme grupos de elementos (unidades) com valores de atributos similares e que a estratificação ideal elimina a variação dentro dos estratos, o que é muito difícil na prática. Laar e Akça (2007, p. 262) complementam que, a fim de gerar estratos homogêneos, as variáveis utilizadas para estratificar a população (índice de sítio, idade, topografia do terreno, tipologia florestal, entre outros) devem estar fortemente relacionadas com a quantidade medida.

Se uma amostra aleatória simples for tomada em cada estrato, então o processo de amostragem será descrito como amostragem estratificada aleatória (COCHRAN, 1977, p. 89).

Segundo Hush *et al.* (1982, p. 171) a amostragem estratificada em inventários florestais possui as seguintes vantagens em relação a amostragem aleatória simples:

- Estimativas separadas das médias e variâncias podem ser obtidas para cada estrato da floresta;
- Para uma mesma intensidade de amostragem, frequentemente a amostragem estratificada produz estimativas mais precisas dos parâmetros da população, se comparada com a amostragem aleatória simples. Isso ocorre se a estratificação estabelecida resultar em uma maior homogeneidade das unidades amostrais dentro de um estrato em relação a população como um todo.

Por outro lado, as desvantagens da estratificação são: o tamanho de cada estrato deve ser conhecido, ou, no mínimo, razoavelmente estimado; e as unidades

amostrais devem ser tomadas em cada um dos estrato, dos quais se queira obter a estimativa (HUSH, MILLER e BEERS, 1982, p. 171).

Péllico Netto e Brena (1997, p. 142) afirmam que ao se estratificar uma população florestal pela primeira vez, pode-se avaliar seu efeito nas estimativas dos estratos por meio de uma análise de variância que mostra se existe ou não diferença significativa entre as médias dos estratos. Esses autores colocam ainda que havendo diferença entre as médias, a amostragem estratificada terá vantagens no que se refere à precisão e custo do inventário se comparada com a amostragem aleatória simples.

O Quadro 9 demonstra as principais estimativas obtidas por meio da amostragem estratificada definidas por Péllico Netto e Brena (1997, p. 135-140).

Estimativa	Fórmula
Média por estrato	$\bar{x}_h = \frac{\sum_{i=1}^{n_h} X_{ih}}{n_h}$
Média estratificada	$\bar{x}_{st} = \frac{\sum_{h=1}^L n_h \cdot \bar{x}_h}{n} = \sum_{h=1}^L w_h \cdot \bar{x}_h$
Variância por estrato	$s_h^2 = \frac{\sum_{i=1}^{n_h} (X_{ih} - \bar{x}_h)^2}{n_h - 1}$
Variância estratificada	$s_{st}^2 = \sum_{h=1}^L w_h \cdot s_h^2$
Variância da média estratificada	$s_{\bar{x}(st)}^2 = \sum_{h=1}^L w_h^2 \cdot \frac{s_h^2}{n_h} \cdot (1 - f_h)$
Erro padrão	$s_{\bar{x}(st)} = \sqrt{s_{\bar{x}(st)}^2}$
Erro de amostragem absoluto	$E_a = \pm t \cdot s_{\bar{x}(st)}$
Erro de amostragem relativo	$E_r = \pm \frac{t \cdot s_{\bar{x}(st)}}{\bar{x}_{st}}$
Intervalo de confiança para a média	$IC \left[ \bar{x}_{st} - t \cdot s_{\bar{x}(st)} \leq \bar{X} \leq \bar{x}_{st} + t \cdot s_{\bar{x}(st)} \right] = P$
Total por estrato	$\hat{X}_h = N_h \cdot \bar{x}_h$
Total da população	$\hat{X} = \sum_{h=1}^L \hat{X}_h = N \cdot \bar{x}_{st}$
Total da população	$\hat{X} = N \cdot \bar{x}$
Intervalo de confiança para o total	$IC \left[ \hat{X} - N \cdot t \cdot s_{\bar{x}(st)} \leq X \leq \hat{X} + N \cdot t \cdot s_{\bar{x}(st)} \right] = P$



Estimativa	Fórmula
Número efetivo de graus de liberdade	$n_e = \frac{\left( \sum_{h=1}^L g_h \cdot s_h^2 \right)^2}{\sum_{h=1}^L g_h^2 \cdot s_h^2} \quad \text{onde: } g_h = \frac{N_h \cdot (N_h - n_h)}{n_h}$
<b>Notação</b> $\bar{x}_h$ = Estimativa da média do estrato $h$ $X_{ih}$ = Variável de interesse da unidade amostral $i$ do estrato $h$ $n_h$ = Quantidade de unidades amostradas no estrato $h$ $\bar{x}_{st}$ = Estimativa da média estratificada $L$ = Quantidade total de estratos $n$ = Quantidade total de unidades amostradas na população $w_h = n_h/n$ = Proporção do estrato $h$ na amostra total $s_h^2$ = Estimativa da variância do estrato $h$ $s_{st}^2$ = Estimativa da variância estratificada $s_{\bar{x}(st)}^2$ = Estimativa da variância da média estratificada $f_h = n_h/N_h$ = Fração amostral do estrato $h$ $N_h$ = Número potencial de unidades do estrato $h$ $s_{\bar{x}(st)}$ = Estimativa do erro padrão estratificado $E_a$ = Erro de amostragem absoluto $t$ = Valor bicaudal da distribuição de Student para $n_e$ graus de liberdade e probabilidade de $1 - P$ $P$ = Probabilidade de confiança $n_e$ = Número efetivo de graus de liberdade $E_r$ = Erro de amostragem relativo $\hat{X}_h$ = Estimativa do total para o estrato $h$ $\hat{X}$ = Estimativa do total da população $N$ = Quantidade total de unidades amostrais da população $g_h$ = Graus de liberdade para o estrato $h$	

QUADRO 9 – ESTIMATIVAS PARA AMOSTRAGEM ESTRATIFICADA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 134-140)

A intensidade de amostragem é obtida da mesma maneira que na amostragem aleatória simples (ver seção 2.1.2), com a particularidade de se utilizar a variância estratificada, conforme demonstrado no Quadro 10 (PÉLLICO NETTO e BRENA, 1997, p. 140).

	Intensidade de amostragem
<b>População finita</b>	$n = \frac{1}{\frac{(LE \cdot \bar{x}_{st})^2}{t^2 \cdot s_{st}^2} + \frac{1}{N}}$
<b>População infinita</b>	$n = \frac{t^2 \cdot s_{st}^2}{(LE \cdot \bar{x}_{st})^2}$

**Notação** $n$  = Intensidade de amostragem $LE$  = Limite do erro de amostragem admitido $\bar{x}_{st}$  = Estimativa da média estratificada $t$  = Valor bicaudal da distribuição de *Student* para  $n_a - 1$  graus de liberdade e probabilidade de  $1 - P$  $n_a$  = Intensidade de amostragem anterior $P$  = Probabilidade de confiança $s_{st}^2$  = Estimativa da variância estratificada $N$  = Quantidade total de unidades amostrais

QUADRO 10 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM ESTRATIFICADA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 140)

O mesmo algoritmo (apresentado na Figura 1), aplicado para amostragem aleatória simples, é utilizado para ajustar a intensidade de amostragem para o caso da amostragem estratificada.

#### 2.1.4 Amostragem sistemática

Para Hush *et al.* (1982, p. 181-182), a amostragem sistemática é um tipo de aplicação das amostragens não aleatórias, em que as unidades de amostragens são amostradas em um intervalo fixo por toda a população. Laar e Akça (2007, p. 300) complementam que as unidades amostrais não são selecionadas aleatoriamente, mas de acordo com um esquema rígido. Esses autores afirmam ainda que a principal vantagem da amostragem sistemática é a simplicidade da sua implementação.

Segundo Köhl *et al.* (2006, p. 93), na amostragem sistemática a população é comumente dividida em uma lista exaustiva de unidades amostrais espaciais, em que  $n$  unidades amostrais são selecionadas dessa lista. Mas, o primeiro passo consiste em selecionar uma unidade amostral aleatoriamente, e então, dessa posição inicial aleatória, uma seleção sistemática das  $n-1$  unidades amostrais restantes é realizada.

Cochran (1977, p. 205) exemplifica a amostragem sistemática da seguinte forma: Supondo que as  $N$  unidades amostrais de uma população estão numeradas de 1 até  $N$  em alguma ordem, para selecionar uma amostra de  $n$  unidades, escolhe-se uma unidade aleatoriamente das primeiras  $k$  unidades e todas as  $k$ -ésimas unidades

depois disso. Por exemplo, se  $k$  é 15 e se a primeira unidade escolhida é a número 13, as unidades subsequentes são as de número 28, 43, 58, e assim por diante. A seleção da primeira unidade determina toda a amostra.

As vantagens da amostragem sistemática são apresentadas por Hush *et al.* (1982, p. 182), conforme o que segue:

- Proporciona uma estimativa confiável do total e da média da população, devido à distribuição uniforme da amostra em toda a população.
- É usualmente executada de forma mais barata e rápida se comparada à amostragem aleatória simples, porque a escolha das unidades amostrais é mecânica, eliminando a necessidade de um processo de seleção aleatório.
- O deslocamento entre as sucessivas unidades amostrais é mais fácil, uma vez que uma direção fixa é preestabelecida, o que resulta em um menor tempo gasto e, conseqüentemente, em um menor custo de amostragem.
- O tamanho da população não precisa ser conhecido, uma vez que as unidades amostrais são escolhidas em um intervalo fixo após um ponto inicial ser estabelecido.
- Cochran (1977, p. 205) complementa que a amostragem sistemática, em geral, é mais precisa do que a amostragem aleatória simples, porque estratifica a população em  $n$  estratos de  $k$  unidades. Por consequência, é esperado que a amostra sistemática seja quase tão precisa quanto a correspondente amostra estratificada com uma unidade por estrato, com a diferença de que na amostragem sistemática, as unidades são selecionadas na mesma posição relativa dentro do estrato, enquanto na amostragem estratificada, a posição das unidades é independente e aleatória.

Laar e Akça (2007, p. 300) citam algumas possíveis desvantagens da amostragem sistemática, quais sejam:

- A amostragem sistemática é um conjunto de unidades amostrais selecionadas de forma não aleatória e, portanto, ela representa uma

amostra de tamanho 1. Em consequência, não é possível estimar a variância da população. Por razões práticas, no entanto, os dados obtidos por amostragem sistemática são analisados com as fórmulas que se aplicam à amostragem aleatória. Isso se justifica se as unidades individuais não estiverem correlacionadas ou nenhuma tendência espacial estiver aparente.

- As unidades amostrais que estão na borda tendem a ser evitadas, portanto as estimativas da média da população podem ser negativamente tendenciosas.
- Para uma única amostra sistemática, fórmulas disponíveis para estimar a variância da média exigem o conhecimento de padrões da distribuição da população.

Segundo Péllico Netto e Brena (1997, p. 160-170), a distribuição sistemática das unidades amostrais pode ser feita com parcelas de área fixa ou faixas e também parcelas de área variável, quando forem usados pontos amostrais ou linhas. Além disso, esses autores classificam a amostragem sistemática em estágio único e dois estágios. O estágio único é caracterizado pela seleção da amostra mediante uma única etapa de amostragem e pode ser realizada por meio de faixas ou parcelas. A amostragem sistemática em dois estágios é realizada em duas etapas de amostragem, cada uma delas com um intervalo de amostragem.

O Quadro 11 demonstra as principais estimativas obtidas por meio da amostragem sistemática definidas por Péllico Netto e Brena (1997, p. 171-175) e separadas de acordo com a classificação proposta por eles.

Estimativa	Fórmula	
	Estágio Único	Dois Estágios
Média	$\bar{x} = \frac{\sum_{i=1}^n X_i}{n}$	$\bar{x} = \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} X_{ij}}{m \cdot n_j}$

Estimativa	Fórmula	
	Estágio Único	Dois Estágios
Variância da média	$s_{\bar{x}}^2 \cong \frac{\sum_{i=1}^n (X_i - \bar{x})^2}{n \cdot (n-1)} \cdot (1-f)$	$s_{\bar{x}}^2 \cong \frac{\sum_{j=1}^m \sum_{i=1}^{n_j} (X_{ij} - X_{(i+1)j})^2}{2 \cdot n \cdot \sum_{j=1}^m (n_j - 1)} \cdot (1-f)$
Erro padrão	$s_{\bar{x}} = \sqrt{s_{\bar{x}}^2}$	
Erro de amostragem absoluto	$E_a = \pm t \cdot s_{\bar{x}}$	
Erro de amostragem relativo	$E_r = \pm \frac{t \cdot s_{\bar{x}}}{\bar{x}}$	
Intervalo de confiança para a média	$IC \left[ \bar{x} - t \cdot s_{\bar{x}} \leq \bar{X} \leq \bar{x} + t \cdot s_{\bar{x}} \right] = P$	
Total da população	$\hat{X} = N \cdot \bar{x}$	
Intervalo de confiança para o total	$IC \left[ \hat{X} - N \cdot t \cdot s_{\bar{x}} \leq X \leq \hat{X} + N \cdot t \cdot s_{\bar{x}} \right] = P$	
<b>Notação</b> $\bar{x}$ = Estimativa da média $X_i$ = Variável de interesse da unidade amostral $i$ $n$ = Quantidade de unidades amostradas $m$ = Quantidade de linhas ou faixas $n_j$ = Quantidade de subunidades dentro da linha ou faixa $j$ $X_{ij}$ = Variável de interesse da unidade amostral $i$ da linha ou faixa $j$ $f = n/N$ = Fração amostral $N$ = Quantidade total de unidades amostrais $s_{\bar{x}}^2$ = Estimativa da variância da média $s_{\bar{x}}$ = Estimativa do erro padrão $E_a$ = Erro de amostragem absoluto $t$ = Valor bicaudal da distribuição de <i>Student</i> para $n-1$ graus de liberdade e probabilidade de $1-P$ $P$ = Probabilidade de confiança $E_r$ = Erro de amostragem relativo $\hat{X}$ = Estimativa do total da população		

QUADRO 11 – ESTIMATIVAS PARA AMOSTRAGEM SISTEMÁTICA  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 171-175)

A intensidade de amostragem na amostragem sistemática, segundo Péllico Netto e Brena (1997, p. 175), é calculada seguindo os mesmos procedimentos usados na amostragem aleatória simples (ver seção 2.1.2). Esses autores destacam ainda que, entretanto, ao se realizar um inventário piloto, sua estrutura amostral deve ser sistemática e preferencialmente ter a mesma estrutura a ser utilizada no inventário definitivo.

### 2.1.5 Amostragem em dois estágios

Segundo Hush *et al.* (1982, p. 176-179) nas amostragens multiestágios, a população consiste de uma lista de unidades amostrais (primeiro estágio), no qual cada uma delas é formada por unidades amostrais menores (segundo estágio), que por sua vez podem ser compostas por unidades amostrais ainda menores (terceiro estágio) e assim por diante. Uma amostra aleatória é obtida para as unidades amostrais do primeiro estágio, uma subamostra é obtida para cada unidade primária, e o procedimento deve continuar até o estágio desejado. Ainda segundo esses autores, a amostragem em dois estágios é uma forma da amostragem multiestágio que está limitada a dois estágios, mas que não está restrita a unidades amostrais de área fixa e pode ser empregada com unidades de área variável, como o método de Bitterlich por exemplo.

A principal vantagem desse processo, segundo Hush *et al.* (1982, p. 176-179), é a redução dos custos da amostragem resultante da concentração de trabalho dentro das unidades primárias amostradas e também a redução dos erros não amostrais devido à supervisão e controle mais efetivos dos trabalhos de campo, facilitados pela concentração das unidades amostrais em compartimentos menores. Cochran (1977, p. 328) cita como vantagem a maior flexibilidade em relação à amostragem de estágio único.

Na opinião de Péllico Netto e Brena (1997, p. 182), a organização estrutural parte da divisão da área total da população em  $N$  unidades primárias, das quais são selecionadas aleatoriamente  $n$  unidades para a amostra. Cada unidade primária é constituída de  $M$  unidades secundárias, que são unidades amostrais definidas para o inventário, das quais  $m$  unidades são aleatoriamente selecionadas para a amostra em cada unidade primária eleita.

Péllico Netto e Brena (1997, p. 182) complementam que a seleção aleatória das unidades primárias e secundárias pode ser sistemática em ambos os estágios ou

pode ser mista quando as unidades primárias são aleatórias e as secundárias são sistemáticas.

O Quadro 12 apresenta as principais estimativas obtidas por meio da amostragem em dois estágios definidas por Péllico Netto e Brena (1997, p. 184-187).

Estimativa	Fórmula
Média da população por subunidade	$\bar{x} = \frac{\sum_{i=1}^n \sum_{j=1}^m X_{ij}}{n \cdot m}$
Média das subunidades por unidade primária	$\bar{x}_i = \sum_{j=1}^m \frac{X_{ij}}{m}$
Variância por subunidade	$s_x^2 = \frac{1}{n \cdot m - 1} \sum_{i=1}^n \sum_{j=1}^m (X_{ij} - \bar{x})^2$
Variância dentro das unidades primárias ou entre as subunidades	$s_d^2 = \frac{\sum_{i=1}^n \sum_{j=1}^m (X_{ij} - \bar{x}_i)^2}{n \cdot (m - 1)}$
Variância entre as unidades primárias	$s_e^2 = \frac{MQ_{entre} - MQ_{dentro}}{m} \quad \text{onde}$ $MQ_{entre} = \frac{\sum_{i=1}^n m \cdot (\bar{x}_i - \bar{x})^2}{n - 1}$ $MQ_{dentro} = s_d^2$
Variância da média	$s_{\bar{x}}^2 = \left( \frac{N - n}{N} \right) \frac{s_e^2}{n} + \left( \frac{M - m}{M} \right) \frac{s_d^2}{n \cdot m}$
Erro padrão	$s_{\bar{x}} = \sqrt{s_{\bar{x}}^2}$
Erro de amostragem absoluto	$E_a = \pm t \cdot s_{\bar{x}}$
Erro de amostragem relativo	$E_r = \pm \frac{t \cdot s_{\bar{x}}}{\bar{x}} \cdot 100$
Intervalo de confiança para a média	$IC \left[ \bar{x} - t \cdot s_{\bar{x}} \leq \bar{X} \leq \bar{x} + t \cdot s_{\bar{x}} \right] = P$
Total da população	$\hat{X} = N \cdot M \cdot \bar{x}$
Intervalo de confiança para o total	$IC \left[ \hat{X} - N \cdot M \cdot t \cdot s_{\bar{x}} \leq X \leq \hat{X} + N \cdot M \cdot t \cdot s_{\bar{x}} \right] = P$
<b>Notação</b> $\bar{x}$ = Estimativa da média da população por subunidade $X_{ij}$ = Variável de interesse da unidade amostral secundária $j$ da unidade amostral primária $i$ $n$ = Quantidade de unidades primárias amostradas $m$ = Quantidade de unidades secundárias amostradas por unidade primária $N$ = Quantidade de unidades primárias da população $M$ = Quantidade de unidades secundárias por unidade primária $\bar{x}_i$ = Estimativa da média das subunidades por unidade primária $s_x^2$ = Estimativa da variância por subunidade $s_d^2$ = Estimativa da variância dentro das unidades primárias ou entre as subunidades $s_e^2$ = Estimativa da variância entre as unidades primárias $MQ_{entre}$ = Quadrado médio entre as unidades primárias	

**Notação**

$MQ_{dentro}$  = Quadrado médio dentro

$s_{\bar{x}}^2$  = Estimativa da variância da média

$s_{\bar{x}}$  = Estimativa do erro padrão

$E_a$  = Erro de amostragem absoluto

$E_r$  = Erro de amostragem relativo

$t$  = Valor bicaudal da distribuição de *Student* para  $n \cdot m - 1$  graus de liberdade e probabilidade de  $1 - P$

$P$  = Probabilidade de confiança

$\hat{X}$  = Estimativa do total da população

QUADRO 12 – ESTIMATIVAS PARA AMOSTRAGEM EM DOIS ESTÁGIOS  
FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 184-187)

A intensidade de amostragem é definida, segundo Péllico Netto e Brena (1997, p. 188-191), para as unidades primárias ( $n$ ) e para as unidades secundárias ( $m$ ), conforme o Quadro 13.

Intensidade de Amostragem		
Unidades Primárias		Unidades Secundárias
<b>População Finita</b> $n = \frac{t^2 \cdot \left( s_e^2 + \frac{s_d^2}{m} \right)}{t^2 \cdot s_{\bar{x}}^2 + \frac{1}{N} \cdot t^2 \cdot \left( s_e^2 + \frac{s_d^2}{M} \right)}$	<b>População Infinita</b> $n = \frac{t^2 \cdot \left( s_e^2 + \frac{s_d^2}{m} \right)}{t^2 \cdot s_{\bar{x}}^2}$	$m = \sqrt{\frac{C_1 \cdot s_d^2}{C_2 \cdot s_e^2}}$
<b>Notação</b> $n$ = Quantidade de unidades primárias amostradas $m$ = Quantidade de unidades secundárias amostradas por unidade primária $N$ = Quantidade de unidades primárias da população $M$ = Quantidade de unidades secundárias por unidade primária $s_d^2$ = Estimativa da variância dentro das unidades primárias ou entre as subunidades $s_e^2$ = Estimativa da variância entre as unidades primárias $s_{\bar{x}}^2$ = Estimativa da variância da média $t$ = Valor bicaudal da distribuição de Student para $n - 1$ graus de liberdade e probabilidade de $1 - P$ $P$ = Probabilidade de confiança $C_1$ = Custo médio de deslocamento $C_2$ = Custo médio de medição		

QUADRO 13 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM DUPLA  
FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 187-191)



### 2.1.6 Amostragem em conglomerados

Segundo Queiroz (2012, p. 233), a amostragem em conglomerados ou grupos é uma variação de qualquer processo de amostragem que, no lugar de utilizar unidades amostrais individuais, usa um grupo ou conglomerado de pequenas subparcelas, que, por sua vez, são denominadas “unidades secundárias”. Péllico Netto e Brena (1997, p. 201) complementam que a amostragem em conglomerado é uma variação da amostragem em dois estágios, onde o segundo estágio é sistematicamente organizado dentro do primeiro estágio de amostragem. Por essa razão, esses autores classificam esse processo como amostragem mista quanto à estrutura organizacional na população amostrada.

Laar e Akça (2007, p. 282-283) afirmam que a amostragem em conglomerados pode trazer benefícios na eficiência do processo para grandes áreas inventariadas, uma vez que ela economiza tempo de deslocamento. Queiroz (2012, p. 234) diz que a amostragem em conglomerado apresenta as seguintes vantagens: oferece melhor controle no trabalho de campo, maior variabilidade da variável de interesse e é adequada em inventários de florestas tropicais de difícil acesso.

Köhl *et al.* (2006, p. 98) citam que os conglomerados podem ter várias formas: quadrados, retângulos ou outras formas mais complexas.

O Quadro 14 demonstra as principais estimativas obtidas por meio da amostragem em conglomerados.

Estimativa	Fórmula
Média da população por subunidade	$\bar{x} = \frac{\sum_{i=1}^n \sum_{j=1}^M X_{ij}}{n \cdot M}$
Média das subunidades por conglomerado	$\bar{x}_i = \sum_{j=1}^M \frac{X_{ij}}{M}$
Variância por subunidade	$s_x^2 = \frac{1}{n \cdot M - 1} \sum_{i=1}^n \sum_{j=1}^M (X_{ij} - \bar{x})^2$
Variância dentro dos conglomerados ou entre as subunidades	$s_d^2 = \frac{\sum_{i=1}^n \sum_{j=1}^m (X_{ij} - \bar{x}_i)^2}{n \cdot (M - 1)}$

Estimativa	Fórmula
Variância entre os conglomerados	$s_e^2 = \frac{MQ_{entre} - MQ_{dentro}}{M} \quad \text{onde}$ $MQ_{entre} = \frac{\sum_{i=1}^n M \cdot (\bar{x}_i - \bar{x})^2}{n-1}$ $MQ_{dentro} = s_d^2$
Coeficiente de correlação intraconglomerados	$r = \frac{s_e^2}{s_e^2 + s_d^2}$
Variância da média	$s_{\bar{x}}^2 = \left( \frac{N-n}{N} \right) \frac{s_e^2}{n} + \frac{s_d^2}{n \cdot m}$
Erro padrão	$s_{\bar{x}} = \sqrt{s_{\bar{x}}^2}$
Erro de amostragem absoluto	$E_a = \pm t \cdot s_{\bar{x}}$
Erro de amostragem relativo	$E_r = \pm \frac{t \cdot s_{\bar{x}}}{\bar{x}} \cdot 100$
Intervalo de confiança para a média	$IC \left[ \bar{x} - t \cdot s_{\bar{x}} \leq \bar{X} \leq \bar{x} + t \cdot s_{\bar{x}} \right] = P$
Total da população	$\hat{X} = N \cdot M \cdot \bar{x}$
Intervalo de confiança para o total	$IC \left[ \hat{X} - N \cdot M \cdot t \cdot s_{\bar{x}} \leq X \leq \hat{X} + N \cdot M \cdot t \cdot s_{\bar{x}} \right] = P$
<b>Notação</b> $\bar{x}$ = Estimativa da média por subunidade $X_{ij}$ = Variável de interesse da subunidade $j$ do conglomerado $i$ $n$ = Quantidade de conglomerados amostrados $N$ = Quantidade de conglomerados da população $M$ = Quantidade de subunidades por conglomerado $\bar{x}$ = Estimativa da média da população por subunidade $\bar{x}_i$ = Estimativa da média das subunidades por conglomerado $s_x^2$ = Estimativa da variância por subunidade $s_d^2$ = Estimativa da variância dentro dos conglomerados ou entre as subunidades $s_e^2$ = Estimativa da variância entre os conglomerados $r$ = Coeficiente de correlação intraconglomerados $MQ_{entre}$ = Quadrado médio entre $MQ_{dentro}$ = Quadrado médio dentro $s_x^2$ = Estimativa da variância total $s_{\bar{x}}^2$ = Estimativa da variância da média $s_{\bar{x}}$ = Estimativa do erro padrão $E_a$ = Erro de amostragem absoluto $E_r$ = Erro de amostragem relativo $t$ = Valor bicaudal da distribuição de <i>Student</i> para $n \cdot M - 1$ graus de liberdade e probabilidade de $1 - P$ $P$ = Probabilidade de confiança $\hat{X}$ = Estimativa do total da população	

QUADRO 14 – ESTIMATIVAS PARA AMOSTRAGEM EM CONGLOMERADOS  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 202-209)

A intensidade de amostragem é definida, segundo Péllico Netto e Brena (1997, p. 188-191), para os conglomerados ( $n$ ) e para as subunidades ( $M$ ), conforme o Quadro 15.

Intensidade de Amostragem	
Conglomerados	Subunidades
$n = \frac{t^2 \cdot s_x^2}{t^2 \cdot s_x^2 \cdot M} \cdot [1 + r \cdot (M - 1)]$	$M = \sqrt{\frac{C_1}{C_2} \cdot \frac{1-r}{r}}$
<p><b>Notação</b></p> <p><math>n</math> = Quantidade de conglomerados amostrados</p> <p><math>N</math> = Quantidade de conglomerados da população</p> <p><math>M</math> = Quantidade de subunidades por conglomerado</p> <p><math>M</math> = Quantidade de unidades secundárias por unidade primária</p> <p><math>s_x^2</math> = Estimativa da variância por subunidade</p> <p><math>s_x^2</math> = Estimativa da variância da média</p> <p><math>r</math> = Coeficiente de correlação intraconglomerados</p> <p><math>t</math> = Valor bicaudal da distribuição de <i>Student</i> para <math>n - 1</math> graus de liberdade e probabilidade de <math>1 - P</math></p> <p><math>P</math> = Probabilidade de confiança</p> <p><math>C_1</math> = Custo médio de deslocamento</p> <p><math>C_2</math> = Custo médio de medição</p>	

QUADRO 15 – INTENSIDADE DE AMOSTRAGEM PARA AMOSTRAGEM EM CONGLOMERADOS  
 FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 210)

### 2.1.7 Amostragem sistemática com múltiplos inícios aleatórios

A amostragem sistemática com múltiplos inícios aleatórios é uma concepção de amostragem em que a estrutura amostral é uma amostragem sistemática em dois estágios, convergindo para um procedimento aleatório. A estrutura de amostragem é similar à da amostragem em conglomerados, podendo-se proceder a uma análise de variância para se estimar os componentes de variância entre as faixas e dentro delas (QUEIROZ, 2012, p. 121). Laar e Akça (2007, p. 301) também afirmam que esse processo de amostragem se assemelha à amostragem por conglomerados e possui média e variâncias amostrais determinadas para cada *cluster* separadamente.

Segundo Péllico Netto e Brena (1997, p. 225) esse processo de amostragem foi proposto por Shiue (1960) que afirma que todas as faixas ou linhas de unidades

devem ser selecionadas aleatoriamente, caracterizando, assim, a denominação de múltiplos inícios aleatórios.

Péllico Netto e Brena (1997, p. 228) citam que a orientação da amostragem, visando minimizar a variância entre linhas ou faixas, é a principal vantagem desse processo para aplicação em inventários florestais.

Segundo Sanquetta *et al.* (2009, p. 147), cada início aleatório é tratado como um conglomerado nesse processo, e as fórmulas utilizadas no processo de amostragem em conglomerados podem ser empregadas. Portanto, o Quadro 14 demonstra as principais estimativas obtidas por meio da amostragem sistemática com múltiplos inícios aleatórios e o Quadro 15 apresenta como se obter a intensidade de amostragem para esse processo.

#### 2.1.8 Amostragem em múltiplas ocasiões

Segundo Queiroz (2012, p. 365), o procedimento para avaliar o caráter dinâmico do crescimento de uma floresta é conhecido com Inventário Florestal Contínuo, que corresponde a efetuar medições repetidas vezes no tempo, ou seja, medições em várias ocasiões.

As amostragens em múltiplas ocasiões, ou ocasiões sucessivas nos inventários florestais, segundo Hush *et al.* (1982, p. 188-189), têm três objetivos principais:

- Estimar as quantidades e as características da floresta presentes no primeiro inventário;
- Estimar as quantidades e as características da floresta presentes no segundo inventário;
- Estimar as mudanças ocorridas na floresta entre os dois inventários.

Ainda, de acordo com Hush *et al.* (1982, p. 188-189), o processo pode ser repetido e, nas ocasiões de todos os subsequentes inventários, o inventário anterior se torna o primeiro e o novo inventário se torna o segundo.

Péllico Netto e Brena (1997, p. 237) definem três grupos de unidades amostrais que são utilizados nos quatro processos de amostragem em múltiplas ocasiões. Cada grupo de unidades amostrais é identificado com uma letra minúscula como segue:

- $(m)$  – Grupo de unidades amostrais permanentes, que é remedido em todas as ocasiões.
- $(u)$  – Grupo de unidades amostrais temporárias, que é medido somente na primeira ocasião.
- $(n)$  – Grupo de unidades novas temporárias, que é medido somente na segunda ocasião.

Nas quatro seções seguintes serão apresentados os processos básicos de inventários florestais contínuos e que foram implementados no *software* educacional discutido nesse trabalho.

### **Amostragens independentes**

Neste processo, as duas ocasiões são executadas independentemente, aplicando-se a amostragem aleatória simples, ou seja, na primeira ocasião a amostra necessária para estimar os parâmetros da população, com a precisão desejada e probabilidade de confiança fixada, é constituída de unidades do grupo  $(u)$ . Na segunda ocasião toma-se uma nova amostra da população, totalmente independente da primeira ocasião, constituída de unidades do grupo  $(n)$ . (PÉLLICO NETTO e BRENA, 1997, p. 237)

O Quadro 16 demonstra as principais estimativas para a primeira ocasião, segunda ocasião e crescimento ou mudança para o processo de amostragem em múltiplas ocasiões independentes.

Estimativa	Fórmula
Média na primeira ocasião	$\bar{x}_u = \frac{\sum_{i=1}^{n_1} X_{ui}}{n_1}$
Variância na primeira ocasião	$s_{x_u}^2 = \frac{\sum_{i=1}^{n_1} (X_{ui} - \bar{x}_u)^2}{n_1 - 1}$
Variância da média na primeira ocasião	$s_{\bar{x}_u}^2 = \frac{s_{x_u}^2}{n_1} \cdot \left(1 - \frac{n_1}{N_1}\right)$
Erro padrão na primeira ocasião	$s_{\bar{x}_u} = \sqrt{s_{\bar{x}_u}^2}$
Erro de amostragem absoluto na primeira ocasião	$E_{au} = \pm t_u \cdot s_{\bar{x}_u}$
Erro de amostragem relativo na primeira ocasião	$E_{ru} = \pm \frac{t_u \cdot s_{\bar{x}_u}}{\bar{x}_u} \cdot 100$
Intervalo de confiança para a média na primeira ocasião	$IC[\bar{x}_u - t_u \cdot s_{\bar{x}_u} \leq \bar{X} \leq \bar{x}_u + t_u \cdot s_{\bar{x}_u}] = P$
Total da população na primeira ocasião	$\hat{X} = N \cdot \bar{x}_u$
Intervalo de confiança para o total na primeira ocasião	$IC[\hat{X} - t_u \cdot N \cdot s_{\bar{x}_u} \leq X \leq \hat{X} + t_u \cdot N \cdot s_{\bar{x}_u}] = P$
Média na segunda ocasião	$\bar{y}_n = \frac{\sum_{h=1}^{n_2} Y_{nh}}{n_2}$
Variância na segunda ocasião	$s_{y_n}^2 = \frac{\sum_{h=1}^{n_2} (Y_{nh} - \bar{y}_n)^2}{n_2 - 1}$
Variância da média na segunda ocasião	$s_{\bar{y}_n}^2 = \frac{s_{y_n}^2}{n_2} \cdot \left(1 - \frac{n_2}{N_2}\right)$
Erro padrão na segunda ocasião	$s_{\bar{y}_n} = \sqrt{s_{\bar{y}_n}^2}$
Erro de amostragem absoluto na segunda ocasião	$E_{an} = \pm t_n \cdot s_{\bar{y}_n}$
Erro de amostragem relativo na segunda ocasião	$E_{rn} = \pm \frac{t_n \cdot s_{\bar{y}_n}}{\bar{y}_n} \cdot 100$
Intervalo de confiança para a média na segunda ocasião	$IC[\bar{y}_n - t_n \cdot s_{\bar{y}_n} \leq \bar{Y} \leq \bar{y}_n + t_n \cdot s_{\bar{y}_n}] = P$
Total da população na segunda ocasião	$\hat{Y} = N \cdot \bar{y}_n$
Intervalo de confiança para o total na segunda ocasião	$IC[\hat{Y} - t_n \cdot N \cdot s_{\bar{y}_n} \leq Y \leq \hat{Y} + t_n \cdot N \cdot s_{\bar{y}_n}] = P$
Crescimento médio	$\bar{d}_i = \bar{y}_n - \bar{x}_u$
Variância da média do crescimento	$s_{\bar{d}_i}^2 = s_{\bar{x}_u}^2 + s_{\bar{y}_n}^2$
Erro padrão do crescimento	$s_{\bar{d}_i} = \sqrt{s_{\bar{d}_i}^2}$
Erro de amostragem absoluto do crescimento	$E_{ac} = \pm t_c \cdot s_{\bar{d}_i}$
Erro de amostragem relativo do crescimento	$E_{rc} = \pm \frac{t_c \cdot s_{\bar{d}_i}}{\bar{d}_i} \cdot 100$
Intervalo de confiança para a média do crescimento	$IC[\bar{d}_i - t_c \cdot s_{\bar{d}_i} \leq \bar{D} \leq \bar{d}_i + t_c \cdot s_{\bar{d}_i}] = P$
Crescimento total	$\hat{D}_i = N \cdot \bar{d}_i$

Estimativa	Fórmula
Intervalo de confiança para o total do crescimento	$IC\left[\hat{D}_i - t_c \cdot N \cdot s_{\bar{d}_i} \leq D \leq \hat{D}_i + t_c \cdot N \cdot s_{\bar{d}_i}\right] = P$
<b>Notação</b> $\bar{x}_u$ = Estimativa da média na primeira ocasião $X_{ui}$ = Variável de interesse $i$ do grupo de unidades temporárias ( $u$ ) $n_1$ = Quantidade de unidades amostrais do grupo de unidades temporárias ( $u$ ) $s_{x_u}^2$ = Estimativa da variância na primeira ocasião $s_{\bar{x}_u}^2$ = Estimativa da variância da média na primeira ocasião $s_{\bar{x}_u}$ = Estimativa do erro padrão na primeira ocasião $E_{au}$ = Erro de amostragem absoluto na primeira ocasião $E_{ru}$ = Erro de amostragem relativo na primeira ocasião $t_u$ = Valor bicaudal da distribuição de <i>Student</i> para $n_1 - 1$ graus de liberdade e probabilidade de $1-P$ $P$ = Probabilidade de confiança $N$ = Quantidade de unidades amostrais da população $N_1$ = Quantidade de unidades amostrais da população na primeira ocasião $\hat{X}$ = Estimativa do total da população na primeira ocasião $\bar{y}_n$ = Estimativa da média na segunda ocasião $Y_{nh}$ = Variável de interesse $h$ do grupo de unidades temporárias ( $n$ ) $n_2$ = Quantidade de unidades amostrais do grupo de unidades temporárias ( $n$ ) $s_{y_n}^2$ = Estimativa da variância na segunda ocasião $s_{\bar{y}_n}^2$ = Estimativa da variância da média na segunda ocasião $s_{\bar{y}_n}$ = Estimativa do erro padrão na segunda ocasião $E_{an}$ = Erro de amostragem absoluto na segunda ocasião $E_{rn}$ = Erro de amostragem relativo na segunda ocasião $t_u$ = Valor bicaudal da distribuição de <i>Student</i> para $n_2 - 1$ graus de liberdade e probabilidade de $1-P$ $N_2$ = Quantidade de unidades amostrais da população na segunda ocasião $\hat{Y}$ = Estimativa do total da população na segunda ocasião $\bar{d}_i$ = Estimativa do crescimento médio $s_{\bar{d}_i}^2$ = Estimativa da variância da média do crescimento $s_{\bar{d}_i}$ = Estimativa do erro padrão do crescimento $E_{ac}$ = Erro de amostragem absoluto do crescimento $E_{rc}$ = Erro de amostragem relativo do crescimento $t_c$ = Valor bicaudal da distribuição de <i>Student</i> para $(n_1 - 1) + (n_2 - 1)$ graus de liberdade e probabilidade de $1 - P$ $\hat{D}_i$ = Estimativa do crescimento total	

QUADRO 16 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES INDEPENDENTES

FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 238-244)

### Amostragem com repetição total

Neste processo, as unidades amostrais são remedidas nas sucessivas ocasiões, ou seja, não são medidas novas unidades amostrais em ocasiões subsequentes. Portanto, todas as unidades amostrais são permanentes e pertencem

ao grupo ( $m$ ). Esse processo também é conhecido como Inventário Florestal Contínuo - IFC (PÉLLICO NETTO e BRENA, 1997, p. 249-250).

Os mesmos estimadores da primeira e segunda ocasião do processo de amostragens independentes, apresentados no Quadro 16, são utilizados nesse processo.

Os estimadores de mudança ou crescimento para o processo de amostragem com repetição total estão apresentados no Quadro 17.

Estimativa	Fórmula
Crescimento médio	$\bar{d}_m = \bar{y}_m - \bar{x}_m$
Variância da média do crescimento	$s_{\bar{d}_m}^2 = s_{\bar{x}_m}^2 + s_{\bar{y}_m}^2 - \frac{2 \cdot s_{xy}}{m}$
Covariância entre as duas ocasiões	$s_{xy} = \frac{\sum_{j=1}^m (X_j - \bar{x}_m) \cdot (Y_j - \bar{y}_m)}{m-1}$
Erro padrão do crescimento	$s_{\bar{d}_m} = \sqrt{s_{\bar{d}_m}^2}$
Erro de amostragem absoluto do crescimento	$E_{ac} = \pm t_c \cdot s_{\bar{d}_m}$
Erro de amostragem relativo do crescimento	$E_{rc} = \pm \frac{t_c \cdot s_{\bar{d}_m}}{\bar{d}_m} \cdot 100$
Intervalo de confiança para a média do crescimento	$IC \left[ \bar{d}_m - t_c \cdot s_{\bar{d}_m} \leq \bar{D} \leq \bar{d}_m + t_c \cdot s_{\bar{d}_m} \right] = P$
Crescimento total	$\hat{D}_m = N \cdot \bar{d}_m$
Intervalo de confiança para o total do crescimento	$IC \left[ \hat{D}_m - t_c \cdot N \cdot s_{\bar{d}_m} \leq D \leq \hat{D}_m + t_c \cdot N \cdot s_{\bar{d}_m} \right] = P$
<b>Notação</b> $\bar{y}_m$ = Estimativa da média na segunda ocasião $\bar{x}_m$ = Estimativa da média na primeira ocasião $\bar{d}_m$ = Estimativa do crescimento médio $s_{\bar{x}_m}^2$ = Estimativa da variância da média na primeira ocasião $s_{\bar{y}_m}^2$ = Estimativa da variância da média na segunda ocasião $s_{xy}$ = Covariância entre as duas ocasiões $s_{\bar{d}_m}^2$ = Estimativa da variância da média do crescimento $s_{\bar{d}_m}$ = Estimativa do erro padrão do crescimento $m$ = Número de unidades amostrais permanentes (igual na primeira e na segunda ocasião) $X_j$ = Variável de interesse $j$ do grupo de unidades permanentes ( $m$ ) da primeira ocasião $Y_j$ = Variável de interesse $j$ do grupo de unidades permanentes ( $m$ ) da segunda ocasião $E_{ac}$ = Erro de amostragem absoluto do crescimento $E_{rc}$ = Erro de amostragem relativo do crescimento $P$ = Probabilidade de confiança $t_c$ = Valor bicaudal da distribuição de <i>Student</i> para $(n_1 - 1) + (n_2 - 1)$ graus de liberdade e probabilidade de $1 - P$	



**Notação** $\hat{D}_m$  = Estimativa do crescimento total**QUADRO 17 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO TOTAL**

FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 252-254)

**Amostragem dupla**

Segundo Hush *et al.* (1982, p. 192), para este tipo de processo de amostragem uma parte das unidades amostrais medidas na primeira ocasião é remedida na segunda ocasião.

Pode-se considerar que a intensidade de amostragem da primeira ocasião, na amostragem dupla, é  $(n_1 = u + m)$ . Dessa amostra obtém-se uma média geral  $(\bar{x}_1)$ , referente às  $(n_1)$  unidades, uma média  $(\bar{x}_u)$  referente à subamostra temporária, e uma média  $(\bar{x}_m)$ , relativa a subamostras permanentes. Na segunda ocasião o inventário é repetido na mesma área, com uma intensidade amostral  $(n_2 = m)$ , tal que  $(n_2)$  seja uma subamostragem de  $(n_1)$ , ou seja,  $n_2 < n_1$ , onde se remede apenas a subamostra permanente  $(m)$ , e se obtém a média  $(\bar{y}_m)$  (PÉLLICO NETTO e BRENA, 1997).

O crescimento do inventário no período, segundo Hush *et al.* (1982, p. 192), é expresso pela diferença entre a média de todo o inventário inicial e a regressão estimada para o segundo inventário. Para Péllico Netto e Brena (1997, p. 262), as estimativas das características de interesse, na segunda ocasião, são obtidas por meio de uma regressão linear, ajustada com base nos dados da subamostra permanente medidos em ambas as ocasiões, e expressa por  $(\bar{y}_r)$ , conforme equação apresentada no Quadro 18.

O Quadro 18 apresenta as principais estimativas para a primeira ocasião, segunda ocasião e crescimento ou mudança para o processo de amostragem dupla em múltiplas ocasiões independentes.

Estimativa	Fórmula
Proporção de unidades temporárias na primeira ocasião	$P_u = \frac{u}{n_1}$
Proporção de unidades permanentes na primeira ocasião	$P_m = \frac{m}{n_1}$

Estimativa	Fórmula
Unidades amostrais medidas na primeira ocasião	$n_1 = u + m$
Média das unidades temporárias na primeira ocasião	$\bar{x}_u = \frac{\sum_{i=1}^u X_{ui}}{u}$
Média das unidades permanentes na primeira ocasião	$\bar{x}_m = \frac{\sum_{j=1}^m X_{mj}}{m}$
Média na primeira ocasião	$\bar{x}_1 = P_u \cdot \bar{x}_u + P_m \cdot \bar{x}_m$
Variância das unidades temporárias na primeira ocasião	$s_{x_u}^2 = \frac{\sum_{i=1}^u (X_{ui} - \bar{x}_u)^2}{u - 1}$
Variância das unidades permanentes na primeira ocasião	$s_{x_m}^2 = \frac{\sum_{j=1}^m (X_{mj} - \bar{x}_m)^2}{m - 1}$
Variância na primeira ocasião	$s_{x_1}^2 = \frac{\sum_{i=1}^{n_1} (X_i - \bar{x}_1)^2}{n_1 - 1}$
Variância da média na primeira ocasião	$s_{\bar{x}_1}^2 = \frac{s_{x_1}^2}{n_1} \cdot \left(1 - \frac{n_1}{N_1}\right)$
Erro padrão na primeira ocasião	$s_{\bar{x}_1} = \sqrt{s_{\bar{x}_1}^2}$
Erro de amostragem absoluto na primeira ocasião	$E_{a1} = \pm t_1 \cdot s_{\bar{x}_1}$
Erro de amostragem relativo na primeira ocasião	$E_{r1} = \pm \frac{t_1 \cdot s_{\bar{x}_1}}{\bar{x}_1} \cdot 100$
Intervalo de confiança para a média na primeira ocasião	$IC \left[ \bar{x}_1 - t_1 \cdot s_{\bar{x}_1} \leq \bar{X} \leq \bar{x}_1 + t_1 \cdot s_{\bar{x}_1} \right] = P$
Total da população na primeira ocasião	$\hat{X} = N \cdot \bar{x}_1$
Intervalo de confiança para o total na primeira ocasião	$IC \left[ \hat{X} - t_1 \cdot N \cdot s_{\bar{x}_1} \leq X \leq \hat{X} + t_1 \cdot N \cdot s_{\bar{x}_1} \right] = P$
Média na segunda ocasião	$\bar{y}_m = \frac{\sum_{j=1}^{n_2} Y_{mj}}{n_2}$
Média na segunda ocasião por regressão	$\bar{y}_r = \bar{y}_m + b \cdot (\bar{x}_1 - \bar{x}_m) \quad \text{Onde:}$ $b = \frac{\sum_{j=1}^m Y_{mj} \cdot X_{mj}}{\sum_{j=1}^m X_{mj}^2}$
Variância na segunda ocasião	$s_{y_m}^2 = \frac{\sum_{j=1}^{n_2} (Y_{mj} - \bar{y}_m)^2}{n_2 - 1}$
Variância da média na segunda ocasião	$s_{\bar{y}_r}^2 = \frac{s_{xy}^2}{n_2} + \frac{s_y^2 - s_{xy}^2}{n_1}$
Variância da regressão	$s_{yx}^2 = \frac{1}{n_2 - 2} \cdot \left[ \sum_{j=1}^m Y_{mj}^2 - \frac{\left( \sum_{j=1}^m Y_{mj} \cdot X_{mj} \right)^2}{\sum_{j=1}^m X_{mj}^2} \right]$

Estimativa	Fórmula
Erro padrão na segunda ocasião	$s_{\bar{y}_r} = \sqrt{s_{\bar{y}_r}^2}$
Erro de amostragem absoluto na segunda ocasião	$E_{am} = \pm t_m \cdot s_{\bar{y}_r}$
Erro de amostragem relativo na segunda ocasião	$E_{rm} = \pm \frac{t_m \cdot s_{\bar{y}_r}}{\bar{y}_r} \cdot 100$
Intervalo de confiança para a média na segunda ocasião	$IC \left[ \bar{y}_r - t_m \cdot s_{\bar{y}_r} \leq \bar{Y} \leq \bar{y}_r + t_m \cdot s_{\bar{y}_r} \right] = P$
Total da população na segunda ocasião	$\hat{Y} = N \cdot \bar{y}_r$
Intervalo de confiança para o total na segunda ocasião	$IC \left[ \hat{Y} - t_m \cdot N \cdot s_{\bar{y}_r} \leq Y \leq \hat{Y} + t_m \cdot N \cdot s_{\bar{y}_r} \right] = P$
Crescimento médio	$\bar{d}_r = \bar{y}_r - \bar{x}_1$
Variância da média do crescimento	$s_{\bar{d}_r}^2 = \frac{(s_y^2 + s_{yx}^2) \cdot \left[ \frac{1+r \cdot (r-2)}{r^2} \right]}{n_1} + \frac{s_{yx}^2}{n_2}$
Coeficiente de correlação	$r = \frac{Cov(xy)}{\sqrt{s_{x_m}^2 \cdot s_{y_m}^2}}$
Erro padrão do crescimento	$s_{\bar{d}_r} = \sqrt{s_{\bar{d}_r}^2}$
Erro de amostragem absoluto do crescimento	$E_{ac} = \pm t_c \cdot s_{\bar{d}_r}$
Erro de amostragem relativo do crescimento	$E_{rc} = \pm \frac{t_c \cdot s_{\bar{d}_r}}{\bar{d}_r} \cdot 100$
Intervalo de confiança para a média do crescimento	$IC \left[ \bar{d}_r - t_c \cdot s_{\bar{d}_r} \leq \bar{D} \leq \bar{d}_r + t_c \cdot s_{\bar{d}_r} \right] = P$
Crescimento total	$\hat{D}_r = N \cdot \bar{d}_r$
Intervalo de confiança para o total do crescimento	$IC \left[ \hat{D}_r - t_c \cdot N \cdot s_{\bar{d}_r} \leq D \leq \hat{D}_r + t_c \cdot N \cdot s_{\bar{d}_r} \right] = P$
<b>Notação</b> $P_u$ = Proporção de unidades temporárias na primeira ocasião $u$ = Quantidade de unidades amostrais do grupo de unidades temporárias $n_1$ = Quantidade de unidades amostrais medidas na primeira ocasião $P_m$ = Proporção de unidades permanentes na primeira ocasião $m$ = Quantidade de unidades amostrais permanentes $\bar{x}_u$ = Estimativa da média das unidades temporárias na primeira ocasião $X_{ui}$ = Variável de interesse $i$ do grupo de unidades temporárias ( $u$ ) da primeira ocasião $\bar{x}_m$ = Estimativa da média das unidades permanentes na primeira ocasião $X_{mj}$ = Variável de interesse $j$ do grupo de unidades permanentes ( $m$ ) da primeira ocasião $\bar{x}_1$ = Estimativa da média na primeira ocasião $s_{x_u}^2$ = Estimativa da variância das unidades temporárias na primeira ocasião $s_{x_m}^2$ = Estimativa da variância das unidades permanentes na primeira ocasião $s_{x_1}^2$ = Estimativa da variância na primeira ocasião $s_{\bar{x}_1}^2$ = Estimativa da variância da média na primeira ocasião $s_{\bar{x}_1}$ = Erro padrão na primeira ocasião $E_{a1}$ = Erro de amostragem absoluto na primeira ocasião $E_{r1}$ = Erro de amostragem relativo na primeira ocasião $t_1$ = Valor bicaudal da distribuição de <i>Student</i> para $n_1 - 1$ graus de liberdade e probabilidade de $1 - P$ $P$ = Probabilidade de confiança $N$ = Quantidade de unidades amostrais da população	

**Notação**

$N_1$	= Quantidade de unidades amostrais da população na primeira ocasião
$\hat{X}$	= Estimativa do total da população na primeira ocasião
$\bar{y}_m$	= Estimativa da média na segunda ocasião
$Y_{mj}$	= Variável de interesse $j$ do grupo de unidades permanentes ( $m$ ) da segunda ocasião
$n_2$	= Quantidade de unidades amostrais do grupo de unidades permanentes ( $m$ )
$\bar{y}_r$	= Estimativa da média na segunda ocasião por regressão
$s_{ym}^2$	= Estimativa da variância na segunda ocasião
$s_{\bar{y}_r}^2$	= Estimativa da variância da média na segunda ocasião
$s_{xy}^2$	= Variância da regressão
$s_{\bar{y}_r}$	= Estimativa do erro padrão na segunda ocasião
$E_{am}$	= Erro de amostragem absoluto na segunda ocasião
$E_{rm}$	= Erro de amostragem relativo na segunda ocasião
$t_m$	= Valor bicaudal da distribuição de Student para $m - 1$ graus de liberdade e probabilidade de $1 - P$
$N_2$	= Quantidade de unidades amostrais da população na segunda ocasião
$\hat{Y}$	= Estimativa do total da população na segunda ocasião
$\bar{d}_r$	= Estimativa do crescimento médio
$s_{\bar{d}_r}^2$	= Estimativa da variância da média do crescimento
$r$	= Coeficiente de correlação
$s_{\bar{d}_r}$	= Estimativa do erro padrão do crescimento
$E_{ac}$	= Erro de amostragem absoluto do crescimento
$E_{rc}$	= Erro de amostragem relativo do crescimento
$t_c$	= Valor bicaudal da distribuição de Student para $(n_1 - 1) + (m - 1)$ graus de liberdade e probabilidade de $1 - P$
$\hat{D}_r$	= Estimativa do crescimento total

QUADRO 18 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM AMOSTRAGEM DUPLA

FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 263-272)

### Amostragem com repetição parcial

Neste tipo de processo de amostragem a segunda ocasião é uma subamostra da primeira ocasião acrescida de novas unidades amostrais (HUSH, MILLER e BEERS, 1982).

Ware e Cunia (1962), citados por Péllico e Brena (1997, p. 283), destacam a versatilidade desse processo em relação à eventual mudança de precisão de uma ocasião para outra, bem como ao tipo de estratificação usado, ou tipo de medição das árvores, que permite ser adaptado. Esses autores afirmam que sua principal vantagem é a forte correlação existente entre os volumes da primeira e da segunda ocasiões, estimadas a partir das unidades permanentes ( $m$ ). Esta correlação permite ajustar uma regressão linear, que é aplicada às unidades temporárias da primeira

ocasião ( $u$ ) e da segunda ocasião ( $n$ ), obtendo-se as estimativas dos volumes dessas subamostras nas ocasiões em que elas não foram medidas. Os mesmos autores concluem que as estimativas são obtidas por meio da totalidade das unidades amostrais, ou seja,  $(u+m+n)$ , ao invés de somente  $(u+m)$  para a primeira ocasião,  $(m+n)$  para a segunda ocasião e apenas  $(m)$  para o crescimento ou mudança.

Os mesmos estimadores da primeira ocasião do processo de amostragem dupla, apresentados no Quadro 18, são utilizados nesse processo.

Os estimadores da segunda ocasião e mudança ou crescimento para o processo de amostragem com repetição parcial estão apresentados no Quadro 19.

Estimativa	Fórmula
Proporção de unidades permanentes na segunda ocasião	$Q_m = \frac{m}{n_2}$
Proporção de unidades novas na segunda ocasião	$Q_n = \frac{n}{n_2}$
Unidades amostrais medidas na segunda ocasião	$n_2 = m + n$
Média das unidades permanentes na segunda ocasião	$\bar{y}_m = \frac{\sum_{i=1}^m Y_{mi}}{m}$
Média das unidades novas da segunda ocasião	$\bar{y}_n = \frac{\sum_{j=1}^n Y_{nj}}{n}$
Média simples da segunda ocasião	$\bar{y}_2 = Q_u \cdot \bar{y}_u + Q_m \cdot \bar{y}_m$
Média corrente da segunda ocasião por regressão	$\bar{y} = a \cdot \bar{x}_u - a \cdot \bar{x}_m + c \cdot \bar{y}_m + (1-c) \cdot \bar{y}_n$ Onde: $a = \frac{m \cdot P_u}{n_2 - P_u \cdot n \cdot r^2} \cdot b_{yx}$ $c = \frac{m}{n_2 - P_u \cdot n \cdot r^2}$ $r = \frac{Cov(xy)}{s_{x_m} \cdot s_{y_m}}$ $b_{yx} = \frac{Cov(xy)}{s_x^2}$
Variância das unidades permanentes na segunda ocasião	$s_{y_m}^2 = \frac{\sum_{j=1}^m (Y_{mj} - \bar{y}_m)^2}{m-1}$
Variância das unidades novas da segunda ocasião	$s_{y_n}^2 = \frac{\sum_{j=1}^n (Y_{nj} - \bar{y}_n)^2}{n-1}$

Estimativa	Fórmula
Variância da segunda ocasião	$s_{y_2}^2 = \frac{\sum_{j=1}^{n_2} (Y_j - \bar{y}_2)^2}{n_2 - 1}$
Variância da média na segunda ocasião	$s_{\bar{y}}^2 = a^2 \cdot s_x^2 \cdot \left( \frac{1}{u} + \frac{1}{m} \right) + c^2 \cdot \frac{s_y^2}{m} + (1-c)^2 \cdot \frac{s_y^2}{m} - 2 \cdot a \cdot c \cdot r \cdot \frac{s_x \cdot s_y}{m}$
Erro padrão na segunda ocasião	$s_{\bar{y}} = \sqrt{s_{\bar{y}}^2}$
Erro de amostragem absoluto na segunda ocasião	$E_{a2} = \pm t_2 \cdot s_{\bar{y}}$
Erro de amostragem relativo na segunda ocasião	$E_{r2} = \pm \frac{t_2 \cdot s_{\bar{y}}}{\bar{y}} \cdot 100$
Intervalo de confiança para a média na segunda ocasião	$IC \left[ \bar{y} - t_2 \cdot s_{\bar{y}} \leq \bar{Y} \leq \bar{y} + t_2 \cdot s_{\bar{y}} \right] = P$
Total da população na segunda ocasião	$\hat{Y} = N \cdot \bar{y}$
Intervalo de confiança para o total na segunda ocasião	$IC \left[ \hat{Y} - t_2 \cdot N \cdot s_{\bar{y}} \leq Y \leq \hat{Y} + t_2 \cdot N \cdot s_{\bar{y}} \right] = P$
Crescimento médio por estimativa direta	$\bar{d}_p = A \cdot \bar{y}_m + (1-A) \cdot \bar{y}_n + B \cdot \bar{x}_m - (1+B) \cdot \bar{x}_u \text{ Onde}$ $A = \frac{m}{n_2 - P_u \cdot n \cdot r^2} + \frac{n \cdot P_m}{n_2 - P_u \cdot n \cdot r^2} \cdot r \cdot \frac{s_{xm}}{s_{ym}}$ $B = \frac{-m \cdot P_u}{n_2 - P_u \cdot n \cdot r^2} \cdot r \cdot \frac{s_{xm}}{s_{ym}} - \frac{n_2 \cdot P_m}{n_2 - P_u \cdot n \cdot r^2}$
Variância da média do crescimento	$s_{\bar{d}_p}^2 = A^2 \cdot \frac{s_y^2}{m} + (1-A)^2 \cdot \frac{s_y^2}{n} + B^2 \cdot \frac{s_x^2}{m} + (1+B)^2 \cdot \frac{s_x^2}{u} + 2 \cdot A \cdot B \cdot r \cdot \frac{s_{xm} \cdot s_{ym}}{m}$
Erro padrão do crescimento	$s_{\bar{d}_p} = \sqrt{s_{\bar{d}_p}^2}$
Erro de amostragem absoluto do crescimento	$E_{ac} = \pm t_c \cdot s_{\bar{d}_p}$
Erro de amostragem relativo do crescimento	$E_{rc} = \pm \frac{t_c \cdot s_{\bar{d}_p}}{\bar{d}_p} \cdot 100$
Intervalo de confiança para a média do crescimento	$IC \left[ \bar{d}_p - t_c \cdot s_{\bar{d}_p} \leq \bar{D} \leq \bar{d}_p + t_c \cdot s_{\bar{d}_p} \right] = P$
Crescimento total	$\hat{D}_p = N \cdot \bar{d}_p$
Intervalo de confiança para o total do crescimento	$IC \left[ \hat{D}_p - t_c \cdot N \cdot s_{\bar{d}_p} \leq D \leq \hat{D}_p + t_c \cdot N \cdot s_{\bar{d}_p} \right] = P$
<b>Notação</b> $Q_m$ = Proporção de unidades permanentes na segunda ocasião $m$ = Quantidade de unidades amostrais permanentes $n_2$ = Quantidade de unidades amostrais medidas na segunda ocasião $Q_n$ = Proporção de unidades novas na segunda ocasião $n$ = Quantidade de unidades novas medidas na segunda ocasião $P_u$ = Proporção de unidades temporárias na primeira ocasião (vide Quadro 18) $P_m$ = Proporção de unidades permanentes na primeira ocasião (vide Quadro 18) $\bar{y}_m$ = Estimativa da média das unidades permanentes na segunda ocasião $Y_{mi}$ = Variável de interesse $i$ do grupo de unidades permanentes ( $m$ ) da segunda ocasião $\bar{y}_n$ = Estimativa da média das unidades novas na segunda ocasião $Y_{nj}$ = Variável de interesse $j$ do grupo de unidades novas ( $n$ ) da segunda ocasião $\bar{y}_2$ = Estimativa da média simples da segunda ocasião $\bar{y}$ = Estimativa da média corrente da segunda ocasião por regressão	

**Notação**

$r$	= Coeficiente de correlação
$b_{yx}$	= Coeficiente angular da regressão
$s_{ym}^2$	= Estimativa da variância das unidades permanentes na segunda ocasião
$s_{yn}^2$	= Estimativa da variância das unidades novas na segunda ocasião
$s_{y2}^2$	= Estimativa da variância na segunda ocasião
$s_{\bar{y}}^2$	= Estimativa da variância da média na segunda ocasião
$s_{\bar{y}}$	= Estimativa do erro padrão na segunda ocasião
$E_{a2}$	= Erro de amostragem absoluto na segunda ocasião
$E_{r2}$	= Erro de amostragem relativo na segunda ocasião
$t_2$	= Valor bicaudal da distribuição de <i>Student</i> para $n_2 - 1$ graus de liberdade e probabilidade de $1 - P$
$\hat{Y}$	= Estimativa do total da população na segunda ocasião
$\bar{d}_p$	= Estimativa do crescimento médio
$s_{\bar{d}_p}^2$	= Estimativa da variância da média do crescimento
$s_{\bar{d}_p}$	= Estimativa do erro padrão do crescimento
$E_{ac}$	= Erro de amostragem absoluto do crescimento
$E_{rc}$	= Erro de amostragem relativo do crescimento
$t_c$	= Valor bicaudal da distribuição de <i>Student</i> para $(n_1 - 1) + (n_2 - 1)$ graus de liberdade e probabilidade de $1 - P$
$\hat{D}_p$	= Estimativa do crescimento total

QUADRO 19 – ESTIMATIVAS PARA AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO PARCIAL

FONTE: ADAPTADO DE PÉLLICO NETTO E BRENA (1997, P. 284-292)

## 2.2 CONSTRUÇÃO DE SOFTWARE E A MODELAGEM ORIENTADA A OBJETOS

O processo de construção de um *software* envolve diversas disciplinas, tais como, gerenciamento de projetos, engenharia de requisitos, projeto, construção e manutenção de *software*. Essas disciplinas podem formar um conjunto de processos, que são apoiados por ferramentas, métodos e modelos, para então formar um conceito maior: a engenharia de *software* (PRESSMAN, 2006, p. 17 e 18). O *IEEE* (*Institute of Electrical and Electronics Engineers*) estabeleceu a sua definição para o conceito de engenharia de *software*: “Engenharia de *software*: (1) aplicação de uma abordagem sistemática, disciplinada e quantificável, para o desenvolvimento, operação e manutenção do *software*; isto é, a aplicação da engenharia ao *software*. (2) O estudo de abordagens como as de (1).” (IEEE, 1993)

Bourque e Fairley (2014) definem um processo de engenharia como um conjunto de atividades inter-relacionadas que transformam uma ou mais entradas em saídas enquanto consomem recursos para completar a transformação. Mais especificamente, esses autores ainda definem que um processo de engenharia de *software* tem como foco as atividades de trabalho realizadas por engenheiros de *software* para desenvolver, manter e operar o *software*, como requisitos, projetos, construções, testes, gestão de configuração e outros processos de engenharia de *software*.

Nos últimos 30 anos modelos de processos que enfatizam a definição, identificação e aplicação detalhada de atividades e tarefas de processo têm sido aplicados pela comunidade de engenharia de *software* com o objetivo de melhorar a qualidade do sistema. Por outro lado, quando aplicados sem adaptações a cada realidade, eles podem aumentar o nível de burocracia e etapas desnecessárias e criar dificuldades para desenvolvedores e clientes (PRESSMAN, 2006, p. 21).

### 2.2.1 Modelos de processos para construção de *software*

Existem vários modelos de processos para construção de *software* que podem ser encontrados na literatura. As próximas seções apresentam uma pequena revisão dos principais modelos.

No desenvolvimento do *software* deste trabalho, nenhum desses modelos de desenvolvimento de *software* é utilizado na íntegra, pois optou-se por utilizar uma abordagem metodológica dupla, conforme descrito na seção 3.1.

#### **Modelo em cascata**

O modelo em cascata é o mais antigo dos modelos de engenharia de *software* e foi criado na década de 70. Pode ser aplicado quando os requisitos de um problema são razoavelmente conhecidos e quando o trabalho flui de maneira linear. Nesse modelo, uma atividade inicia somente após o término da atividade anterior.



A principal vantagem desse modelo é a simplicidade, e a principal desvantagem é a dificuldade na introdução de modificações ao longo do projeto.

A Figura 2 ilustra esse modelo e os seus 5 passos: comunicação, planejamento, modelagem, construção e implantação.

No primeiro passo (comunicação) o analista entende os requisitos e acorda os principais objetivos do projeto com o cliente. Na fase de planejamento é gerado o cronograma, com uma estimativa de esforço e custo para o projeto. Os modelos de análise e projeto são gerados na terceira etapa desse processo, por meio do uso de uma linguagem de modelagem, como por exemplo, a *UML (Unified Modeling Language)*. Na quarta etapa (construção) o *software* é codificado em uma ou mais linguagens de programação e testado de forma automática ou manual. A implantação é a etapa na qual o *software* é entregue e/ou instalado. No modelo em cascata, somente ao final da última etapa o cliente pode utilizar o *software* em questão.

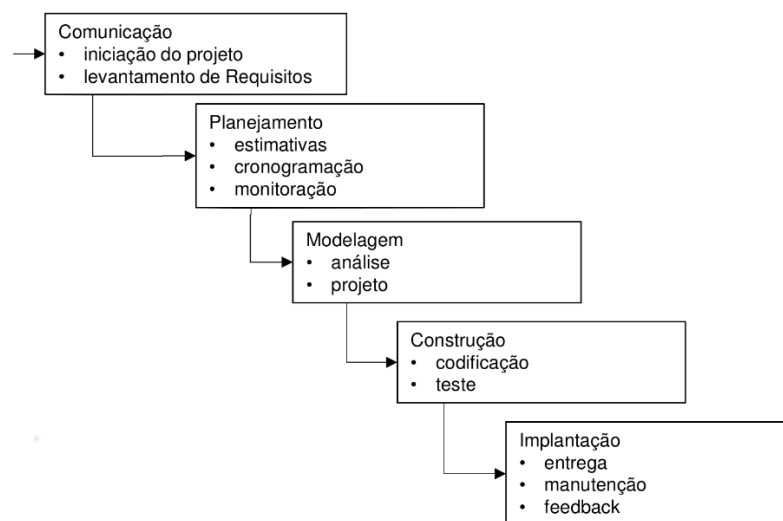


FIGURA 2 – MODELO EM CASCATA DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE  
FONTE: PRESSMAN (2006)

### Modelo incremental

O modelo incremental utiliza o modelo em cascata de maneira iterativa. Conforme demonstrado na Figura 3, cada incremento passa por todas as etapas do modelo em cascata. Normalmente, quando se aplica esse tipo de modelo, um conjunto de requisitos é entregue a cada incremento completo, ou seja, esse modelo apresenta

um produto operacional a cada incremento, até que o produto completo seja produzido.

Esse modelo é mais flexível que o modelo em cascata, pois permite que funcionalidades elencadas sejam entregues antes do fim do projeto para o cliente. Outra característica desse modelo é que, se houver disponibilidade de recursos, os incrementos podem ocorrer simultaneamente, como observado na Figura 3.

Uma desvantagem desse modelo é a complexidade da sua gestão, pois permite que os incrementos ocorram de forma paralela e algumas atividades ocorrerão ao mesmo tempo. Por exemplo, em um determinado projeto a primeira iteração pode estar na fase de construção e a segunda iteração na fase de comunicação, demandando maior habilidade do gerente de projetos e maior maturidade da equipe que está desenvolvendo o projeto para priorizar as tarefas.

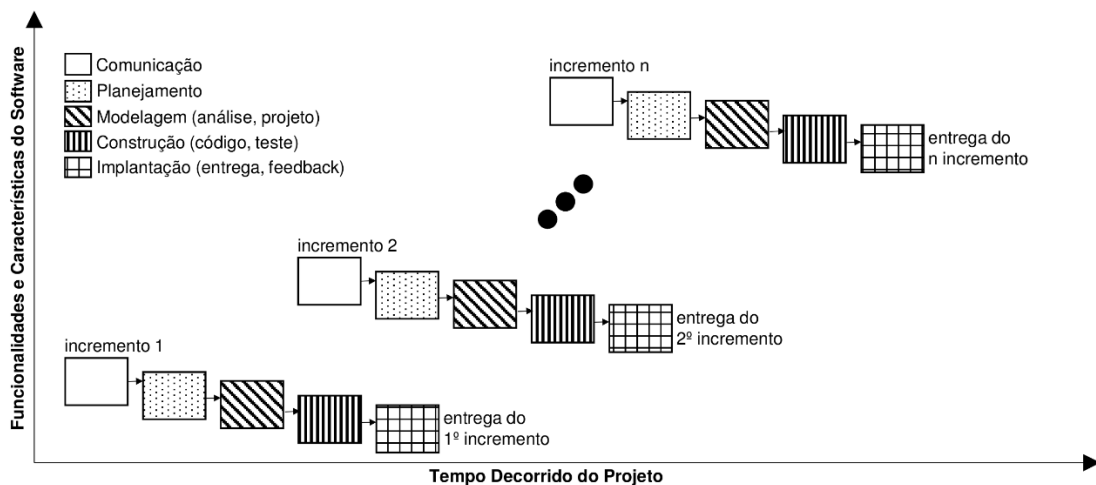


FIGURA 3 – MODELO INCREMENTAL DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE  
FONTE: PRESSMAN (2006)

### Modelo Evolucionário: Prototipagem

Os modelos evolucionários pressupõem que o *software*, como todo sistema complexo, evolui com o passar do tempo (GILB e FINZI, 1988) citado por (PRESSMAN, 2006). Esses modelos também são incrementais, mas diferentemente do modelo incremental nem sempre entregam um produto funcional ao usuário no final de cada ciclo. Por esse motivo são mais flexíveis que os modelos anteriores,

permitindo construir um produto que tenha seus requisitos modificados e incrementados ao longo do tempo.

Um exemplo típico desse modelo é a prototipagem que tem por característica marcante uma modelagem e projeto rápidos para validar com o cliente os requisitos do sistema. Algumas vezes pode-se optar por um protótipo não funcional, em que o principal objetivo é validar questões de layout das telas e formas de navegação da aplicação. Pressman (2006, p. 42) cita que, em alguns casos, o desenvolvedor pode querer avaliar a eficiência de algum algoritmo como a prototipagem. No caso desse trabalho a prototipagem foi utilizada para avaliar a assertividade e eficiência de uma biblioteca de programação (ver seção 4.1). Esse autor, afirma ainda que, a prototipagem é mais comumente usada como uma técnica que pode ser implementada dentro do contexto de qualquer modelo de processo.

Bourque e Fairley (2014) definem prototipagem como um meio para validar a interpretação que o engenheiro de *software* tem dos requisitos do *software*, bem como para elucidar novos requisitos.

As principais vantagens dessa abordagem são: os desenvolvedores conseguem produzir rapidamente um resultado para os clientes; elucidam os requisitos com mais facilidade; e por outro lado, o cliente pode ter resultados mais rápidos com o projeto. Mas, também podem-se citar desvantagens, como um protótipo ser implantado no lugar de um sistema real, causando a operação de um *software* de má qualidade, sem requisitos de escalabilidade, por exemplo. Ou seja, um *software* que funciona bem para um usuário, mas que tem problemas de desempenho quando o número de usuários cresce. Esse modelo está apresentado na Figura 4.

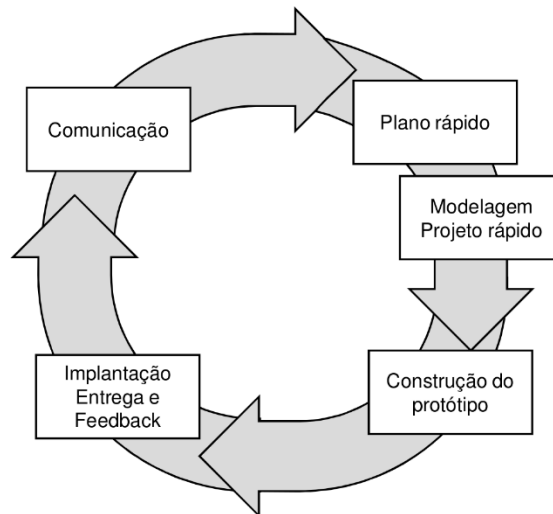


FIGURA 4 – MODELO DE PROTOTIPAÇÃO DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE  
 FONTE: (PRESSMAN, 2006)

### Modelo evolucionário: espiral

Segundo Pressman (2006, p. 44-45), o modelo espiral proposto originalmente por (BOEHM, 1988) é um modelo evolucionário de processo de *software* que combina a natureza iterativa da prototipagem com os aspectos controlados e sistemáticos do modelo em cascata. Podem-se construir rapidamente, com esse modelo, versões de *softwares* cada vez mais completas. Esse modelo aplica-se ao desenvolvimento de *software* de grande porte.

A Figura 5 apresenta o modelo de processo em espiral. A cada volta completa em torno do circuito tem-se uma entrega de um produto. Mas, o produto entregue pode ser uma especificação de um *software*, um protótipo, uma versão do *software* em funcionamento ou uma modificação em um *software* já entregue. Pode-se perceber que esse modelo é bastante flexível; por outro lado, exige uma grande maturidade de gerenciamento para que o produto final seja entregue no prazo e esforço definidos, uma vez que é necessário revisar o plano de projeto a cada volta na espiral.

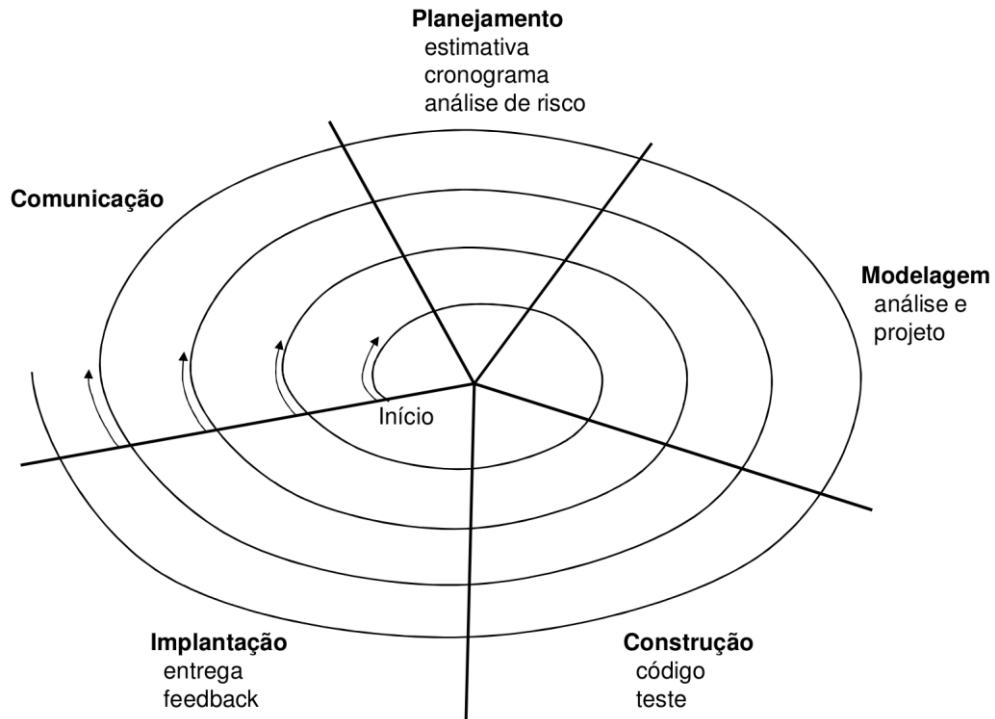


FIGURA 5 – MODELO ESPIRAL DE PROCESSO PARA CONSTRUÇÃO DE SOFTWARE  
 FONTE: (PRESSMAN, 2006)

### Modelo unificado

Na década de 90 alguns autores de processos e métodos se uniram para propor um modelo único de processo. Jacobson, Booch e Rumbaugh (1999) apresentaram o “processo unificado” (*UP – Unified Process*) baseados nos modelos de processos tradicionais (apresentados nas seções anteriores) e no paradigma orientados à objetos.

O processo unificado é dividido em 4 fases: concepção, elaboração, construção e transição. As fases são divididas em iterações, que, por sua vez, passam por seis disciplinas: modelagem de negócios, requisitos, análise e projeto, implementação, testes e implantação. As disciplinas de suporte (gerenciamento de configuração de alteração, gerenciamento de projetos e ambiente) apoiam todo o processo. A Figura 6 ilustra as fases e fluxos de trabalho do processo unificado. Para cada fluxo de trabalho tem-se um histograma que ilustra o esforço ao longo do projeto e suas fases.

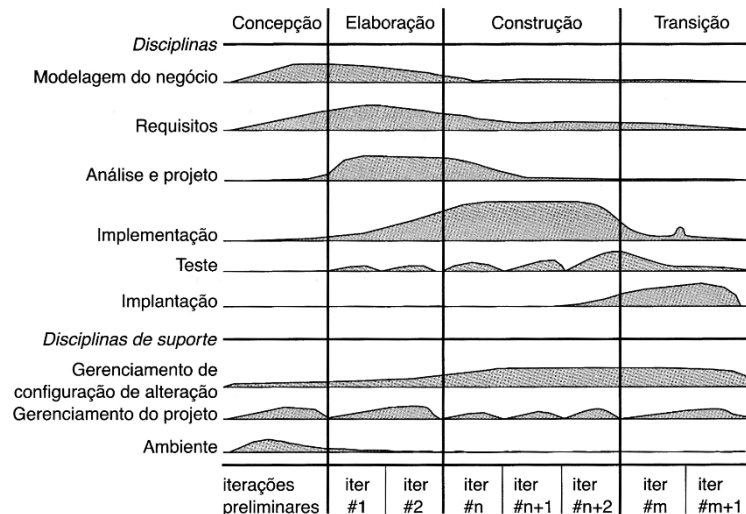


FIGURA 6 – PROCESSO UNIFICADO PARA CONSTRUÇÃO DE SOFTWARE  
 FONTE: (BOOCH, RUMBAUGH E JACOBSON, 2012, P. 490)

Segundo Pressman (2006) o processo unificado é um processo de construção de *software* orientado por casos de uso, centrado na arquitetura, iterativo e incremental, projetado como um conjunto de métodos e ferramentas UML (Unified Modeling Language). Portanto, além do processo, com todos os conjuntos de fases, atividades e melhores práticas, Jacobson, Booch e Rumbaugh (1999) propuseram uma linguagem para modelagem de *software* orientada a objetos reunindo as melhores práticas existentes conhecidas. A seção 2.2.2 descreve os principais conceitos da modelagem orientada a objetos tendo como base a UML.

### Modelo ágil

Na última década o modelo ágil de processo para construção de *software* vem sendo adotado pelas empresas. Mais especificamente no Brasil, isso tem ocorrido nos últimos 5 anos. Essa mudança vem acontecendo, pois os modelos tradicionais de desenvolvimento, como os apresentados nos tópicos anteriores dessa seção, na maioria dos projetos não respondiam bem a mudanças e a projetos de curta duração.

Os modelos de processos tradicionais são guiados por um ciclo de vida que especifica formalmente as tarefas que serão desenvolvidas em cada fase e os papéis relacionados a essas tarefas (analista de sistemas, programadores, testadores, etc.) para os indivíduos que irão desempenhá-las. Esses processos, além do programa

funcionando, produzem uma série de documentos que descrevem o projeto do *software*. Em contrapartida, o modelo ágil lida com a imprevisibilidade dos projetos baseando-se nas pessoas e sua criatividade no lugar de processos definidos. Esse modelo é caracterizado por ciclos curtos de desenvolvimento iterativos direcionados aos requisitos do produto, períodos de reflexão e introspecção, decisões colaborativas, incorporação rápida de mudanças e integração contínua das alterações no sistema durante o desenvolvimento. A entrega de cada ciclo do desenvolvimento é um pedaço do *software* funcionando para o cliente. O modelo ágil desencoraja a documentação por trás do código. Além disso, as estratégias iterativas desse modelo são mais bem suportadas pelas tecnologias orientadas a objetos (NERUR, MAHAPATRA e MANGALARAJ, 2005).

O modelo ágil para desenvolvimento de *software* surgiu a partir do manifesto ágil (BECK *et al.*, 2001) que se baseia em quatro premissas e doze princípios. Podem-se citar alguns métodos de desenvolvimento de *software* que seguem os princípios do manifesto ágil: “*extreme programming*” (XP), “*scrum*” e “*feature-driven development*” (FDD). Mais detalhes sobre esses métodos podem ser encontrados em Pressaman (2006, p. 63-73), Poppendieck e Cusumano (2012).

### 2.2.2 Modelagem orientada a objetos

Nas ciências exatas, mais especificamente nas engenharias, recorrem-se a modelos para representar edifícios, automóveis, motores ou florestas, por exemplo. Cada área do conhecimento possui seus próprios modelos para representar suas realidades. Um edifício pode ser representado por um conjunto de plantas baixas. A eficiência de um motor pode ser representada por um modelo matemático. A quantidade de madeira de uma floresta pode ser representada por um modelo matemático volumétrico.

Os modelos são representações simplificadas de alguma realidade (BOOCH, RUMBAUGH e JACOBSON, 2012). A simplificação é importante nos modelos, pois não é possível representar o sistema ou a realidade modelada completamente em um único modelo. Por exemplo, em uma floresta recorremos a um modelo matemático para representar a quantidade de madeira e a outro modelo para representar o crescimento da floresta. Na mesma floresta existem vários pontos de vista que podem ser modelados, cada um com seu modelo: quantidade de biomassa, volume de madeira, distribuição das alturas, distribuição das espécies, tamanho da floresta em área, densidade, e tantos outros. Não é possível termos todas as informações de uma floresta no mesmo modelo. Isso ocorre em outras áreas - na engenharia civil, por exemplo -, com seus modelos para os diferentes pontos de vista de um edifício: projeto arquitetônico, estrutural, elétrico, hidráulico, entre outros.

Lima (2011) afirma que os melhores modelos expressam a realidade, isto é, quanto maior a complexidade do produto a ser entregue, maior a necessidade de boas técnicas de modelagem, com um conjunto de visões (pontos de vista ou perspectivas) suficientes para garantir sua compreensão.

Mais especificamente para a engenharia de *software*, Bourque e Fairley (2014) afirmam que um modelo é uma abstração ou simplificação de um componente de *software*. Eles acrescentam que, a fim de responder a questões específicas sobre como o *software* deve operar, os modelos são construídos para representar objetos do mundo real e seus comportamentos.

As linguagens de programação orientadas a objetos ganharam popularidade no início da década de 80. No mesmo período surgiram as primeiras linguagens de modelagem orientadas a objetos. Segundo Booch, Rumbaugh e Jacobson (2012) a quantidade de métodos e processos orientados a objetos aumentou de pouco mais de 10 para mais de 50 durante o período de 1989 a 1994. Foi em 1999 que Jacobson, Booch e Rumbaugh (1999) propuseram uma linguagem unificada para modelagem de projetos de *software* orientado a objetos, conhecida como UML, que até os dias atuais serve de padrão para modelagem de sistemas nesse paradigma.



Segundo Booch, Rumbaugh e Jacobson (2012, p. 12), existem dois métodos principais para modelagens de *software*: o da perspectiva de um algoritmo e o da perspectiva da orientação a objetos. Ainda, segundo esses autores, a visão tradicional adota a perspectiva de algoritmo, na qual o principal bloco de construção é o procedimento ou função, que podem ser decompostos para diminuir ou aumentar a complexidade. Por outro lado, eles afirmam que a visão contemporânea adota a perspectiva orientada a objetos, em que o principal bloco de construção é o objeto ou a classe, no qual um objeto é uma representação virtual de uma entidade real e a classe é a descrição desse objeto.

Nas próximas seções serão apresentados os conceitos de classes e objetos, e como representá-los com a UML. O objetivo dessas seções é fornecer os elementos do diagrama de classes, que é um modelo estrutural de um *software* e foi utilizado na seção 4.1 a fim de modelar a biblioteca de programação para processos de amostragem em inventário florestal.

### **Classes e objetos**

Para Booch, Rumbaugh e Jacobson (2012, p. 53), uma classe é uma descrição de um conjunto de objetos que compartilham os mesmos atributos, operações, relacionamentos e semântica. Eles concluem tal afirmação dizendo que se devem utilizar as classes para capturar o vocabulário do sistema que está sendo desenvolvido, e pode-se usar classes para representar itens de *software*, de *hardware* e até itens que são puramente conceituais.

Lima (2011) diz que uma classe é a definição dos atributos e das ações de um tipo de objeto, que descreve um conjunto de objetos individuais em qualquer contexto.

Para Barnes e Kölling (2009, p. 2), os objetos são criados a partir de classes e que elas descrevem o tipo do objeto. Afirmam também que os objetos representam instâncias individuais das classes.

Pode-se afirmar que o conceito de classe é fundamental para modelagem orientada a objetos. Uma classe representa a descrição de uma entidade do mundo real. Pode-se afirmar que uma classe é uma abstração da descrição dessa entidade para um *software*. A classe descreve duas características fundamentais de uma entidade do mundo real: os seus atributos e os seus comportamentos. Para um *software* essas duas características são suficientes para manipular a representação de uma entidade do mundo real no sistema. Os atributos representam as relações, informações ou estados de uma entidade que podem ser armazenadas. Os comportamentos são as ações que podem ser realizadas com essas informações (atributos), com informações de outras entidades relacionadas, ou, simplesmente, com informações fornecidas para essa ação. Comportamento, ação, método, operação ou função são as maneiras como os autores designam a parte dinâmica de uma classe; portanto, pode-se considerar que são sinônimos.

Um objeto é uma entidade do mundo real que é representado no *software* por uma classe. Pode-se afirmar que um conjunto de objetos do mesmo tipo é representado pela mesma classe.

Por exemplo, em um sistema de mensuração florestal, teremos provavelmente as seguintes classes (que são entidades do mundo real): *Arvore*<sup>1</sup> e *Floresta*. Nesse sistema existirá apenas uma classe *Arvore*, mas vários objetos do tipo dessa classe, que representarão para o sistema cada indivíduo de uma floresta. Essa classe (*Arvore*) pode possuir os seguintes atributos: *especie*, *altura*, *DAP* (Diâmetro na Altura do Peito), coordenadas geográficas e as medidas de cada uma de suas seções. Pode-se ainda definir o comportamento, ou as ações que essa classe de objetos pode realizar nas suas informações ou nas informações relacionadas a ela: obter cada uma das informações, calcular o volume por Smailian ou pelo fator de forma.

---

<sup>1</sup> O nome das classes, atributos e operações, por padrão, nas linguagens de programação, não são acentuados, não são separados por espaços para nomes compostos e não possuem caracteres especiais, como cedilhas, por exemplo. Por esse motivo, a partir deste ponto no texto, esses nomes aparecerão em itálico.

A UML tem a capacidade de representar essa classe por meio de um modelo chamado “diagrama de classes” apresentado na Figura 7. Pode-se observar que a classe é representada por um retângulo dividido em três regiões:

- Na parte superior é colocado o nome da classe em negrito (nesse caso, *Arvore*).
- Na parte central são colocados os atributos da classe (nesse caso: *especie*, *altura*, *dap*, *coordenadas* e *secoes*). Na frente do nome dos atributos existe o caractere “-” que representa o nível de visibilidade do atributo (ver Quadro 20). Após o nome de cada atributo lê-se o caractere “:” e então uma palavra, como por exemplo “String” para *especie*. Essa palavra define o tipo do atributo em questão. Nesse caso “String” é uma sequência de caracteres (letras, números e caracteres especiais). Significa que o atributo *especie* pode armazenar uma sequência de letras, que para essa classe representa a espécie de uma Árvore.
- Na parte inferior são colocadas as operações ou métodos da classe (nesse caso: *obterEspecie*, *obterAltura*, *obterDap*, *obterCoordenadas*, *modificarCoordenadas*, *calcularVolumePorSmalian* e *calcularVolumePeloFatorDeForma*). Antes do nome de cada operação existe o caractere “+” que define o nível de visibilidade da operação. Após o nome da operação, entre parênteses, encontra-se o tipo de dado que é esperado como parâmetro para essa operação. O tipo do retorno da operação está após o caractere “:”. A palavra *void* indica que essa operação não tem retorno. Por exemplo, no caso da operação *modificarCoordenadas*, a classe foi programada para alterar as coordenadas do objeto árvore, mas não é necessário que se retorne um valor. Por outro lado, as operações *calcularVolumePorSmalian* e *calcularVolumePeloFatorDeForma* retornam um tipo de dado numérico real, que é o resultado do cálculo do volume pelo método escolhido.

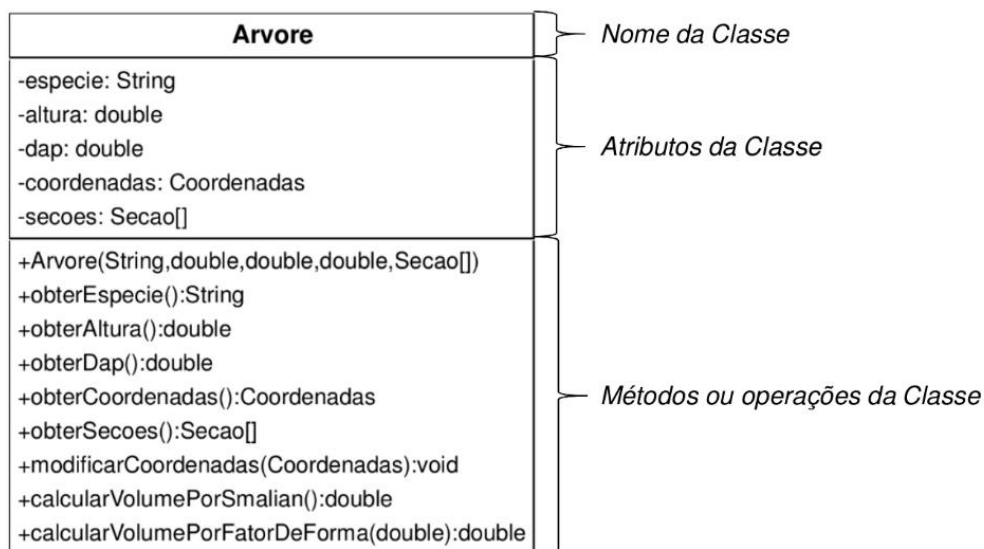


FIGURA 7 – EXEMPLO DE CLASSE QUE REPRESENTA UMA ÁRVORE  
 FONTE: O AUTOR (2014)

Em um *software* orientado a objetos, o programador escreve o código das classes para depois utilizá-las a fim de representar cada objeto do sistema em memória. Então, no caso do sistema de mensuração florestal, a classe *Arvore* seria utilizada para representar cada árvore no sistema. Por exemplo, a árvore da espécie *Pinus elliottii*, de altura 15,3m, dap 8,3 cm, coordenadas (25°26'54.0"S;49°14'27.8"W), e com cada uma das seções medidas seria uma instância ou objeto da classe *Arvore*. Esse objeto existiria no sistema como uma estrutura de dados (atributos) e teria um comportamento específico determinado pelos métodos da sua classe. Cada árvore do sistema seria um objeto específico, mas todos seriam da classe *Arvore*.

Para criar uma instância a partir de uma classe, utiliza-se uma operação ou método especial da classe chamado “construtor”. Esse método tem o mesmo nome da classe. No exemplo da Figura 7, tem-se o construtor *Arvore()*. Entre parêntesis está especificado o tipo dos argumentos esperados para a criação de uma instância de árvore, que, neste caso, são as informações de uma árvore. Poder-se-ia utilizar os dados apresentados no parágrafo anterior para criar uma instância de *Arvore*, cuja espécie fosse *Pinus elliottii*. No exemplo da Figura 8, o construtor para um retângulo espera dois argumentos do tipo *double* (numérico real, ver Quadro 21) que representam a largura e altura do retângulo.

As entidades do mundo real representadas pelas classes, nem sempre são concretas. Por exemplo, pode-se representar um retângulo por meio de uma classe *Retangulo*. Essa classe pode ter como atributos a largura e altura do retângulo e como operações modificar largura, modificar altura e calcular área.

Na Figura 8 está representada a classe *Retangulo* pela UML. A operação *calculaArea* resulta em um tipo *double* (numérico de ponto flutuante) que representa a área do retângulo.

Retangulo
-largura: double -altura: double
+Retangulo(double,double) +modificarAltura(double):void +modificarLargura(double):void +calculaArea():double

FIGURA 8 – EXEMPLO DE CLASSE QUE REPRESENTA UM RETÂNGULO  
FONTE: O AUTOR (2014)

### Visibilidade de membros

Uma das características dos objetos, em uma linguagem de programação orientada a objetos, é a proteção do estado interno do objeto. Os objetos podem se comunicar por meio das suas operações ou métodos. Na maioria dos casos, os atributos de um determinado objeto devem ficar internos a ele. A única forma de acessarmos esses atributos, nesse caso, seria por meio das operações ou métodos.

Para que isso seja possível, define-se na classe, para cada atributo ou operação um nível de visibilidade, que informa se determinado membro do objeto (atributo ou operação) pode ser acessado por outro objeto.

Booch, Rumbaugh e Jacobson (2012, p. 135) afirmam que a visibilidade é um dos mais importantes detalhes que se pode especificar para os atributos e operações. Afirmam ainda que a visibilidade de uma característica específica (atributos ou operações) determina se ela pode ser utilizada por outras classes. Segundo esses autores, na UML, pode-se determinar qualquer um dos quatro níveis de visibilidade.

Segundo Jandl Junior (2007, p. 77) a acessibilidade (ou visibilidade) é um aspecto de extrema importância na programação OO (Orientada a Objetos), pois permite que o programador restrinja certos elementos da classe. Ele afirma também que é com a restrição de acesso que se implementa, em grande parte, o encapsulamento de informações ou implementações dentro das classes.

Boch, Rumbauch e Jacobson (2012) definiram quatro níveis de visibilidade para a UML, que estão definidos no Quadro 20.

Especificador	Símbolo	Descrição
Público	+	Qualquer instância de classe no sistema pode utilizar o membro (atributo ou operação) referenciado por esse especificador. É o menos restritivo.
Protegido	#	Qualquer instância da classe que seja descendente (ver subseção “Relacionamento entre classes”) pode utilizar o membro (atributo ou operação) referenciado por esse especificador.
Privado	-	Pode ser utilizado somente pela própria classe. É o mais restritivo.
Pacote	~	Somente instâncias de classe dentro do mesmo pacote pode utilizar o membro referenciado por esse especificador

QUADRO 20 – ESPECIFICADORES DE ACESSO DE MEMBROS DE CLASSE  
 FONTE: ADAPTADO DE (BOOCH, RUMBAUGH E JACOBSON, 2012, P. 136)

Nos exemplos da Figura 7 e da Figura 8 foram utilizados os especificadores público (+) e privado (-). Os atributos foram definidos como privados e as operações foram definidos como pública para preservar o estado interno dos objetos e apresentar uma interface pública de acesso. Por motivo de simplificação, na modelagem da biblioteca de programação foram definidos os especificadores de visibilidade da mesma maneira.

### **Tipos de dados**

Uma classe é definida pelo conjunto de seus atributos e operações. Cada atributo de uma classe é de um determinado tipo (que pode ser a definição de outra classe). Então, de uma maneira simplista, pode-se afirmar que, quando se define uma classe, um novo tipo de dado é criado. No exemplo da Figura 8, definiu-se um novo tipo (*Retangulo*) que armazena em sua estrutura duas informações (a largura e a

altura). Essas informações armazenadas também são de um determinado tipo, neste caso, inteiro.

Jandl Junior (2007, p. 31) define um tipo de dado como um conjunto particular de valores que podem ser representados dentro de um *software* e quando tais definições fazem parte da linguagem, é dito que constituem um tipo primitivo ou tipo predefinido. Nesse trabalho foi utilizada a linguagem Java, que, segundo Jandl Junior (2007, p. 31), possui oito tipos primitivos, conforme apresentado no Quadro 21.

<b>Categoria</b>	<b>Tipo</b>	<b>Intervalo</b>
Inteiro	byte	de -128 a +127
	short	de -32.768 a +32.767
	int	de -2.147.483.648 a +2.147.483.647
	long	de -9.223.372.036.854.775.808 a +9.223.372.036.854.775.807
Real	float	aproximadamente de -3,4E+38 a +3,4E+38
	double	aproximadamente de -1,7E+308 a +1,7E+308
Caractere	char	de \u0000 a \uFFFF
Lógico	boolean	false e true

QUADRO 21 – TIPOS DE DADOS PRIMITIVOS NA LINGUAGEM JAVA  
FONTE: ADAPTADO DE JANDL JUNIOR (2007, P. 31)

## Relacionamentos entre classes

As classes e, conseqüentemente, suas instâncias, normalmente não são utilizadas de forma isolada. Em um *software* orientado a objetos as classes se relacionam de diferentes maneiras.

Para ilustrar os três tipos de relacionamentos, o exemplo do sistema de mensuração florestal será ampliado. Além da classe *Arvore*, serão incluídas as classes: *Floresta*, *Araucaria*, *Pinus*, *Relatorio* e *Poligono*.

Segundo Booch, Rumbaugh e Jacobson (2012, p. 68-69), existem três tipos de relacionamentos especialmente importantes: dependência, associação e generalização.

## Relacionamentos entre classes: Dependência

A dependência representa um relacionamento de utilização entre as classes. Na UML esse relacionamento é representado por uma linha tracejada. Uma classe pode ter uma relação de dependência com a outra quando utiliza a classe como

argumento ou retorno em uma de suas operações, ou então, quando internamente uma de suas operações utiliza a classe referenciada. (BOOCH, RUMBAUGH e JACOBSON, 2012, p. 70-71)

No exemplo da Figura 9 a classe *Floresta* possui a operação *gerarRelatorio* que retorna um tipo *Relatorio*, ou seja uma instância da classe *Relatorio*, por isso a relação de dependência está representada pela seta pontilhada. A direção da seta indica que a classe *Floresta* depende da classe *Relatorio*.

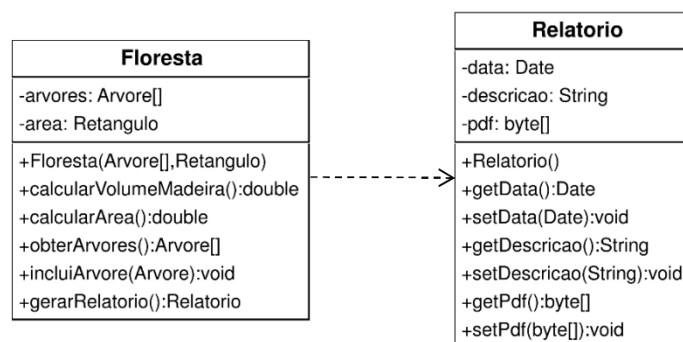


FIGURA 9 – EXEMPLO DE RELACIONAMENTO DE DEPENDÊNCIA ENTRE CLASSES  
FONTE: O AUTOR (2014)

### Relacionamentos entre classes: Associações

As associações são relacionamentos estruturais entre as instâncias das classes. Segundo Booch, Rumbaugh e Jacobson (2012, p. 72) uma associação especifica que um objeto está conectada com um ou mais objetos.

As relações de associação ocorrem quando um dos atributos de uma classe representa uma instância de algum objeto de outra classe, ou em casos específicos da mesma classe.

De forma geral, existem três tipos básicos de associação: associação simples, agregação e composição.

A associação simples é utilizada quando existe uma relação estrutural simples entre dois objetos. Por exemplo, em um sistema de controle acadêmico, um professor pode lecionar uma ou mais disciplinas. Mas as duas entidades, professor e disciplina são independentes. Em um sistema de mensuração florestal, de forma simplificada,



pode-se ter a área de uma floresta representada por um polígono georreferenciado, onde existirá uma relação de associação simples entre a classe *Floresta* e a classe *Poligono*. Esses dois exemplos estão representados respectivamente na Figura 10 (A) e (B).

Outro detalhe importante das associações diz respeito à cardinalidade. Na Figura 10(A), abaixo da seta, está escrito “0..\*”, indicando que um *Professor* pode lecionar de zero até *n* disciplinas. Na Figura 10(B) está escrito “1” abaixo da seta, indicando que uma *Floresta* está associada a um único *Poligono*, o qual representa a sua área.

A palavra acima da seta indica o nome do atributo na classe. O caractere “-” na frente dessa palavra indica que esse atributo é privado à classe (ver Quadro 20). Na Figura 10(A) a palavra *disciplinasLecionadas* refere-se ao atributo da classe *Professor*. O mesmo ocorre com o atributo *area* para a classe *Floresta* na Figura 10(B).

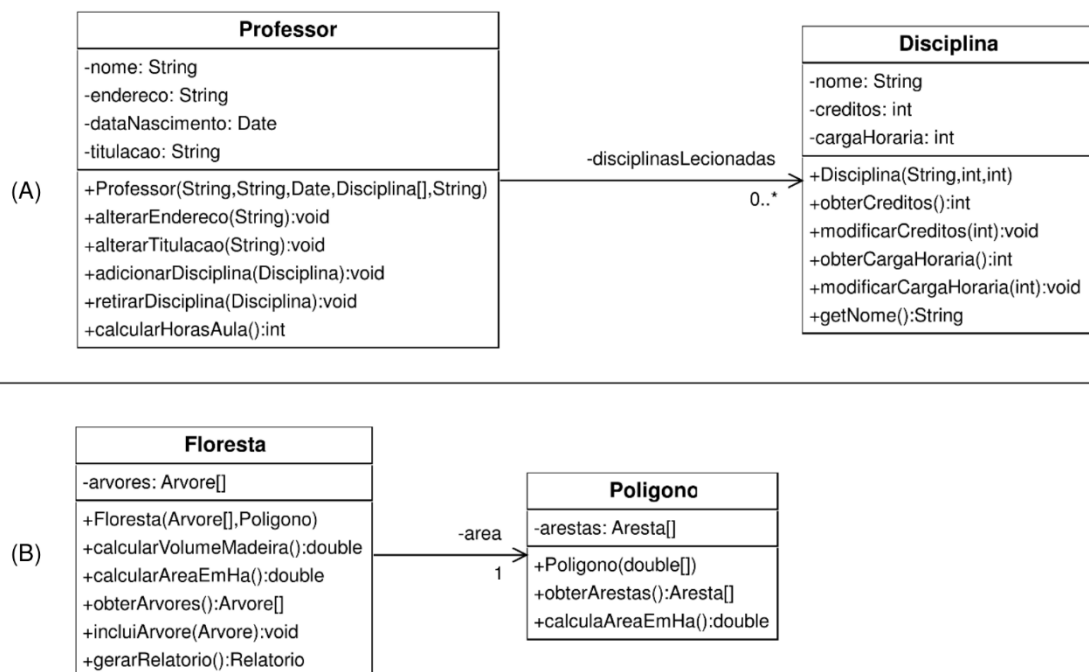


FIGURA 10 – EXEMPLOS DE ASSOCIAÇÃO SIMPLES ENTRE CLASSES  
FONTE: O AUTOR (2014)

A seta indica a direção da navegação entre os objetos. Na Figura 10(A), a partir do objeto da classe *Professor* pode-se obter a lista de disciplinas. Na Figura 10(B), a partir do objeto da classe *Floresta* pode-se obter o *Poligono* que representa a área dessa floresta.

A agregação é utilizada quando existe uma relação parte/todo entre as classes, cuja parte, contudo, não dependa do todo para existir. Por exemplo, em um sistema de mensuração florestal, pode-se ter uma entidade representando uma amostra de árvores. A classe *Amostra* é formada por árvores, mas as árvores continuam existindo se a *Amostra* for desfeita. Essa relação é representada por um diamante vazado próximo à classe agregadora. A Figura 11 demonstra a relação de agregação entre as classes *Amostra* e *Arvore*.

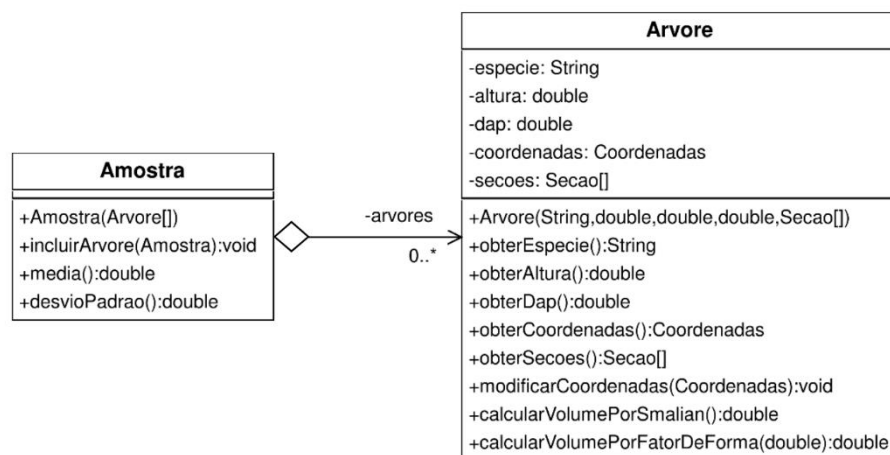


FIGURA 11 – EXEMPLO DE AGREGAÇÃO ENTRE CLASSES  
FONTE: O AUTOR (2014)

A composição, assim como a agregação, também reflete uma relação parte/todo entre as classes, mas é uma relação mais forte do que a anterior. Nesse caso a parte forma o todo e não pode existir isoladamente. Por exemplo, no caso de um sistema de mensuração florestal, pode-se dizer que uma floresta é composta de árvores. Se a floresta deixar de existir, as árvores também deixarão. Essa relação é representada por um diamante cheio próximo à classe agregadora. A Figura 12 demonstra a relação de composição entre as classes *Floresta* e *Arvore*.

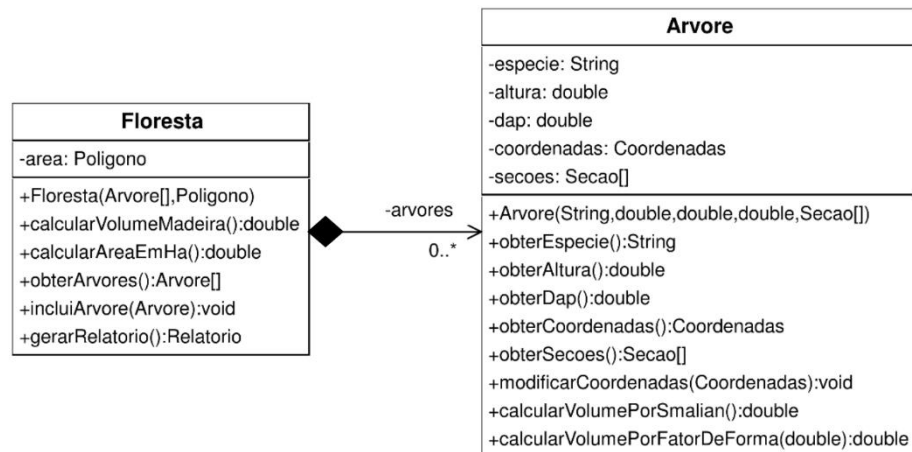


FIGURA 12 – EXEMPLO DE COMPOSIÇÃO ENTRE CLASSES  
 FONTE: O AUTOR(2014)

### Relacionamentos entre classes: Generalizações

Esse tipo de relacionamento também é conhecido como “herança” ou “especialização”. Neste caso uma classe herda as características (atributos e operações) de outra classe. Desse modo, a classe filha (a que herdou) pode se comportar como sua ascendente. Por outro lado, podem-se atribuir operações ou comportamentos específicos para a classe filha.

A generalização ou herança é um mecanismo muito importante para o paradigma de orientação a objetos. Em qualquer linguagem desse tipo teremos esse mecanismo como ponto fundamental dos seus conceitos.

Booch, Rumbaugh e Jacobson (2012, p. 71) definem generalização como um relacionamento entre itens gerais (chamados superclasses ou classes-mãe) e tipos mais específicos desses itens (chamados subclasses ou classes-filhas). Esses autores complementam ainda que muitas vezes as generalizações são chamadas relacionamentos “é um tipo de”. Para Lima (2011, p. 23), uma subclasse herda atributos e comportamento de sua classe imediatamente superior – a superclasse, e também possui atributos e operações específicos dela mesma. Segundo Jandl Junior (2007, p. 108), a herança é um mecanismo que possibilita a uma classe usar campos e métodos definidos em outra classe, o que significa o compartilhamento de membros entre classes.

Por exemplo, em um sistema de mensuração florestal, pode-se ter uma espécie especial de árvore que demande atributos e operações específicos. A árvore da espécie *Araucária angustifolia* pode ser um exemplo. A classe *Araucaria* pode ser uma especialização de uma classe *Arvore*, com atributos adicionais (*sexo* e *producaoPinhaoAnualEmKg*) e operações adicionais: *obterSexo* e *obterProducaoPinhaoAnualEmKg*, conforme demonstrado na Figura 13.

A relação de herança é indicada em um diagrama de classes por meio de uma flecha vazada, cuja direção aponta para a superclasse, conforme observado na Figura 13.

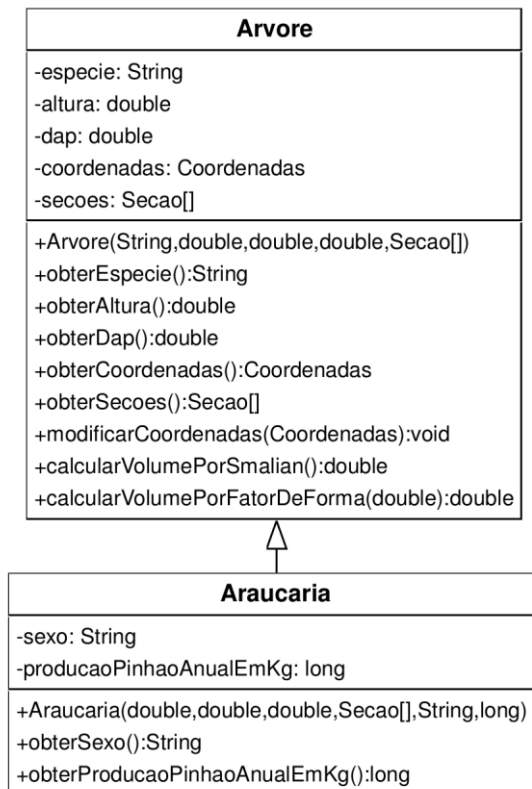


FIGURA 13 – EXEMPLO DE GENERALIZAÇÃO ENTRE CLASSES  
FONTE: O AUTOR (2014)

Neste caso (Figura 13), diz-se que a classe *Araucaria* é do tipo *Arvore*. Quando se cria um objeto a partir da classe *Araucaria*, ele terá como atributos: *especie*, *altura*, *dap*, *coordenadas*, *secoes*, *sexo* e *producaoPinhaoAnualEmKg*. As operações disponíveis nesse objeto do tipo *Araucaria* serão o conjunto das operações

da classe *Arvore* e da classe *Araucaria*. O construtor é uma exceção, pois esses não são herdados; portanto, o único construtor disponível para um objeto do tipo *Araucaria* é o de sua própria classe.

### **Classes abstratas**

Uma classe é abstrata quando não se pode instanciá-la, ou seja, não se pode criar um objeto a partir de uma classe abstrata, ou, conforme afirmam Booch, Rumbaugh e Jacobson (2012, p. 138), não se podem apresentar instâncias diretas. Esse conceito é bastante útil, quando não se pode implementar certas operações de forma genérica.

Por exemplo, para o caso de um sistema de mensuração florestal, pode-se considerar que a classe *Arvore* é abstrata, pois não existe um objeto árvore, mas sim um objeto árvore de uma determinada espécie. Pode-se então, projetar a classe *Arvore* como uma classe abstrata. Por outro lado, somente com essa classe não é possível criar instâncias de árvores para o sistema, tornando-a inútil. Então se deve criar uma relação de herança com essa classe, ou seja, classes de árvores de espécies específicas e com comportamentos específicos, como no exemplo da Figura 13 para a *Araucaria*.

Com esse conceito também é possível criar operações em classes abstratas que não terão implementação. Essas operações serão implementadas nas subclasses da classe abstrata.

No caso de um sistema de mensuração florestal, a operação que calcula o volume da árvore poderia ser abstrata, significando que sua implementação e a forma de obtenção da estimativa do volume poderiam variar de espécie para espécie. Isso dá ao programador a liberdade de escolher o método de cubagem mais adequado a cada espécie. Esse exemplo é demonstrado no diagrama de classes da Figura 14.

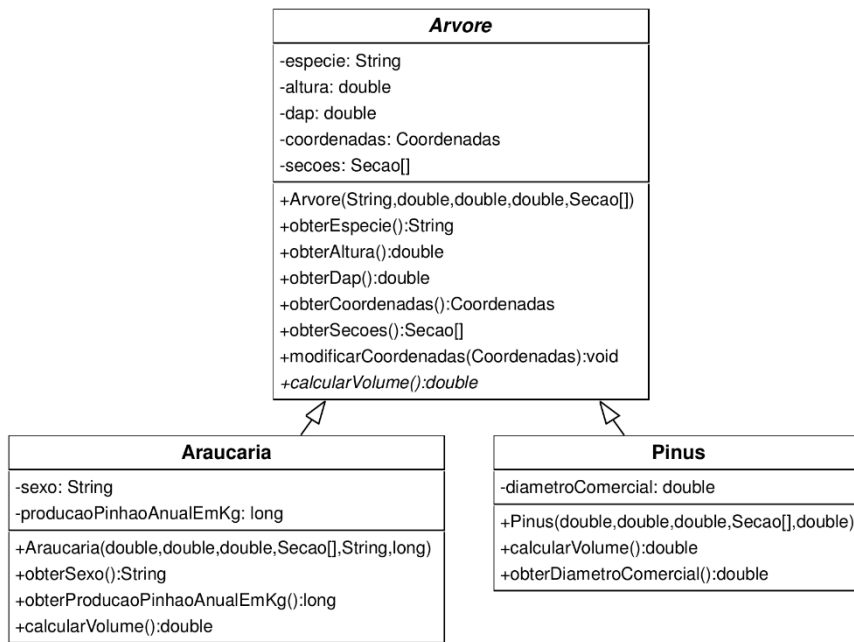


FIGURA 14 – EXEMPLO DE CLASSE ABSTRATA  
FONTE: O AUTOR (2014)

Em um diagrama de classes da UML pode-se identificar uma classe e uma operação abstrata quando o nome está em itálico (caso da classe *Arvore* e da operação *calcularVolume* na mesma classe na Figura 14). Observa-se que o método *calcularVolume* se repete nas subclasses *Araucaria* e *Pinus*. Isso indica que essa operação tem implementações particulares em cada espécie de árvore. Pelo diagrama da Figura 14 pode-se afirmar que um objeto da classe *Araucaria* é uma *Arvore*, assim como um objeto da classe *Pinus*, mas cada um calcula o volume à sua maneira e possui atributos específicos, como *diametroComercial* no caso de *Pinus*.

### Classes Aninhadas

Segundo Jandl Junior (2007, p. 124), classes aninhadas ou internas são definidas como membros da outra classe. Barnes e Kölling (2009, p. 309) complementam que as instâncias (ou objetos) da classe interna são anexadas às instâncias da classe que as incluem – elas só podem existir conjuntamente com uma instância que as incluem. Elas existem conceitualmente dentro da instância que as incluem.

Por exemplo, no caso de um sistema de mensuração florestal, para se aplicar qualquer método de cubagem (HUBER, SMALIAN, NEWTON, HOHENADL, entre outros), é necessária a medição de diâmetros ou circunferências nas várias alturas, ou seja, devem-se seccionar as árvores em toras (MACHADO e FIGUEIREDO FILHO, 2003, p. 224). Portanto, uma classe *Arvore* desse *software* teria as seções das árvores nas suas instâncias. Para representar esse caso poderíamos modelar a classe *Secao* como uma classe interna da classe *Arvore*, uma vez que as instâncias das seções seriam altamente dependentes do objeto árvore. Esse exemplo de modelagem pode ser observado na Figura 15.

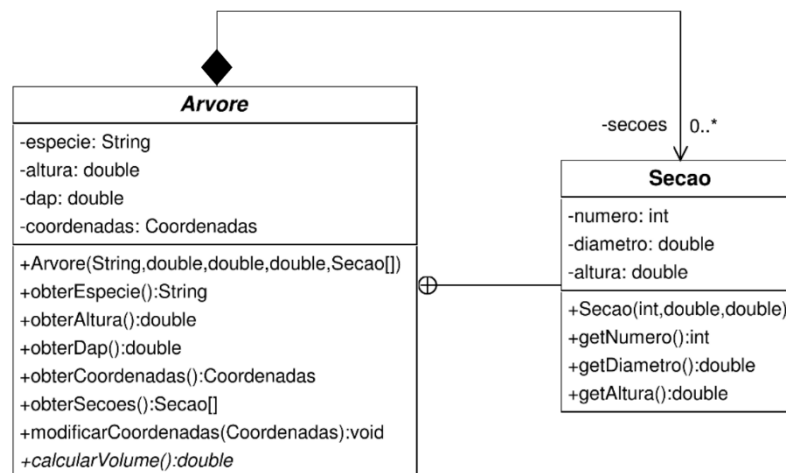


FIGURA 15 – EXEMPLO DE CLASSE ANINHADA  
FONTE: O AUTOR (2014)

Existem duas relações da classe *Arvore* com a classe *Secao* indicada no diagrama da Figura 15: A circunferência vazada com uma cruz (ao lado da classe *Arvore*) indica que a classe *Secao* é aninhada da classe *Arvore*. O diamante cheio (acima da classe *Arvore*) apontando para a classe *Secao* indica uma relação de composição, ou seja, uma árvore é composta por seções.

### Classes genéricas

Segundo Jandl Junior (2007, p. 135) os genéricos são mecanismos para a criação de tipos parametrizados, ou seja, permitem a definição de uma classe ou método para funcionar com uma variedade de tipos. Barnes e Kölling (2009, p. 81)

descrevem que classes genéricas exigem um parâmetro de tipo; em comparação com outras classes, não definem um tipo único, mas potencialmente vários tipos. Boch *et al.* (2012, p. 144) chamam classe genérica de classe *template*, e a descrevem como um modelo de elemento que pode ser parametrizado.

Por exemplo, no caso de um sistema de mensuração florestal, uma classe genérica poderia ser utilizada para definir uma amostra, pois cada amostra refere-se a uma população de um tipo de indivíduos, que pode ser: espécie de árvore, parcela, semente, entre outras. Segundo Péllico Netto e Brena (1997) a amostra pode ser definida como uma parte da população, constituída de indivíduos que apresentam características semelhantes.

No exemplo da Figura 16 está apresentada uma classe genérica *AmostraGenerica*. Sabe-se que é genérica, pois, após o nome da classe, entre os caracteres “<” e “>” está definido o tipo genérico “T”. Significa que a classe *AmostraGenerica* pode ser parametrizada com qualquer tipo. Supondo que, nesse sistema, fosse necessário criar amostras separadas por espécie de árvores, seriam criadas relações de herança com a classe *AmostraGenerica*, e as classes específicas de amostras para cada espécie, no caso desse exemplo: *AmostraAraucaria* e *AmostraPinus*. A indicação do tipo parametrizado das classes *AmostraAraucaria* e *AmostraPinus* está demonstrada na relação de dependência entre as classes *Araucaria* e *Pinus*, respectivamente.

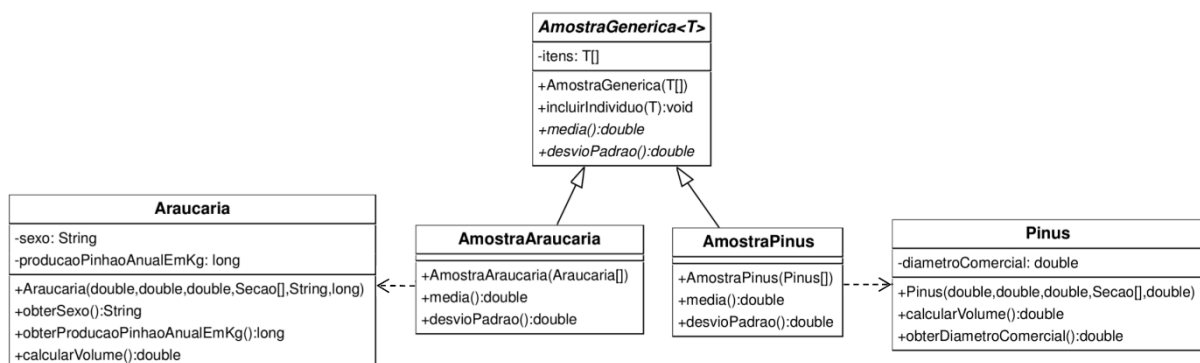


FIGURA 16 – EXEMPLO CLASSE GENÉRICA  
FONTE: O AUTOR (2014)



Além de genérica, a classe *AmostraGenerica*, no exemplo da Figura 16, é abstrata, assim como os métodos *media* e *desvioPadrao*. Esses métodos estão implementados nas classes *AmostraAraucaria* e *AmostraPinus*.

### 2.3 GERAÇÃO ALEATÓRIA DE DADOS COM BASE NO MÉTODO NÃO PARAMÉTRICO NÚCLEO ESTIMADOR VARIÁVEL

A geração aleatória de dados é um tema bastante discutido na Matemática, na Estatística e na Ciência da Computação. O uso de dados aleatórios é aplicado em várias áreas, tais como: criptografia de dados, simulações de ambientes naturais, finanças, jogos eletrônicos, entre outras. Segundo Gentle (2003, p. 1), nas engenharias e nas ciências naturais, simulações são usadas extensivamente nos estudos dos processos físicos e biológicos. Ainda segundo esse autor outro uso comum de números aleatórios ocorre na criptografia de dados. Haar (2014) afirma que números randômicos são úteis para uma variedade de propósitos, tais como: geração de dados para chaves encriptadas, simulações, modelagem de fenômenos complexos e para selecionar amostras randômicas de conjuntos grandes de dados.

Para os inventários florestais a geração de dados aleatórios pode ser útil, por exemplo, quando se utilizam processos computadorizados para sorteio das unidades amostrais que devem participar de uma amostra. Péllico Netto e Brena (1997, p. 107) e Sanquetta *et al.* (2009, p. 84) afirmam que, no caso da amostragem aleatória simples (processo de amostragem a partir do qual derivam todos os procedimentos de amostragem), a seleção das unidades amostrais pressupõe que cada unidade amostral deve ser livre de qualquer escolha e totalmente independente da seleção das demais unidades da amostra.

Segundo Haar (2014) existem duas abordagens principais para a geração de números aleatórios utilizando um computador: pseudogeradores de números

aleatórios (*PRNG – Pseudo Random Number Generation*) e geradores de números aleatórios verdadeiros (*TRNG – True Random Number Generation*).

No caso da abordagem *PRNG*, um computador é utilizado para a geração de números aleatórios de forma determinística, por meio de um algoritmo que espera um ou mais números (sementes) como entrada e gera como saída uma sequência de números aleatórios uniformemente distribuídos entre zero e 1. A entrada esperada por esses algoritmos é chamada de semente, que pode ser qualquer número inteiro disponível. Por exemplo, pode-se utilizar como semente o número do CPF de uma pessoa, a sua data de nascimento, o número de matrícula de um aluno, a data atual do sistema, entre outros. Essa é a abordagem mais comum, pois nenhum computador, por um método puramente computacional é capaz de gerar números completamente aleatórios (GENTLE, 2003, p. 2). Isso se deve à natureza determinística da máquina, ou seja, sem um conjunto de instruções preestabelecidas não há funcionamento.

Vários métodos e algoritmos de geração de números aleatórios pela abordagem *PRNG* foram desenvolvidos e estão descritos na literatura, como por exemplo: “Middle-Square Method” (VON NEUMANN, 1951), Wichmann-Hill (WICHMANN e HILL, 1982), “Middle-Product-Method” (PAUL e BALMER, 1993) e Mersenne-Twister (MATSUMOTO e NISHIMURA, 1998). Os métodos congruentes também podem ser destacados e são revisados na obra de Gentle (2003, p. 5-56).

A geração de números aleatórios verdadeiros é mais difícil, pois exige que algum processo físico ou químico seja integrado a um computador. Gentle (2003, p. 3) cita alguns processos físicos que podem ser utilizados para a geração de números aleatórios reais: Davis *et al.* (1994) descreveu um método de utilização da aleatoriedade da turbulência dos *drives* de discos. A Toshiba produziu um produto comercial, “*Random Master*”, que utilizava o ruído térmico em um semicondutor para gerar sequências uniformes aleatórias.

Por outro lado, Gentle (2003, p. 3) afirma que a maioria das aplicações em estatística, engenharia e ciências naturais podem utilizar a abordagem

pseudoaleatória (PRNG), em que as amostras se assemelham às aleatórias e são baseadas em uma distribuição de probabilidade conhecida.

A distribuição uniforme (equação 2), com intervalo unitário (0,1) é utilizada pela maioria dos geradores de números pseudoaleatórios existentes nas linguagens de programação e nos *softwares* estatísticos.

$$f(x) = \begin{cases} 1 & \text{se } 0 < x < 1; \\ 0 & \text{c.c} \end{cases} \quad (2)$$

Segundo Gentle (2003, p. 5) a distribuição uniforme é conveniente porque existem técnicas simples para transformar as amostras aleatórias obtidas dessa distribuição em amostras de outras distribuições de interesse. Uma das técnicas descritas por esse autor é o método da inversa da função distribuição acumulada (fda) que está descrito na próxima seção.

### 2.3.1 Método da inversa da função distribuição acumulada (fda)

O método da inversa da fda é uma técnica simples para transformar as amostras aleatórias obtidas de uma distribuição uniforme para qualquer outra distribuição de interesse. Ou seja, por meio desse método é possível gerar dados pseudoaleatórios a partir de uma distribuição conhecida.

Esse método, descrito por Devroye (1986, p. 28), é definido por dois passos simples:

- Passo 1: Gerar uma variável uniforme  $U[0,1]$
- Passo 2: Retornar  $x \leftarrow F^{-1}(U)$

Onde  $U[0,1]$  é uma distribuição uniforme no intervalo de 0 a 1 e  $F^{-1}(U)$  é a inversa da função distribuição acumulada, que tem como parâmetro os dados gerados a partir da distribuição uniforme.

Para definir o método, Devroye (1986, p. 28) descreve um teorema que é apresentado na Figura 17. Esse autor afirma que o teorema pode ser utilizado para

gerar dados pseudoaleatórios com uma função de distribuição contínua  $F$ , desde que  $F^{-1}$  seja explicitamente conhecida.

**Teorema.**

Seja  $F$  uma função de distribuição de probabilidade em  $\mathbb{R}$  com a inversa generalizada definida por  $F^{-1}(u) = \inf \{x : F(x) = u, 0 \leq u \leq 1\}$

1. Se  $U$  é uma variável aleatória uniformemente distribuída em  $[0,1]$ , então  $F^{-1}(U)$  possui uma função de distribuição  $F$ .
2. Se  $X$  tem uma função de distribuição  $F$ , então  $F(X)$  está uniformemente distribuída em  $[0,1]$ ;

**Prova do Teorema.**

1. A primeira declaração pode ser provada pelo fato de que para todo  $x \in \mathbb{R}$ ,  

$$P(F^{-1}(U) \leq x) = P(\inf \{y : F(y) = U\} \leq x)$$

$$= P(U \leq F(x)) = F(x)$$
2. A segunda declaração pode ser provada pelo fato de que todo  $u \in [0,1]$ ,  

$$P(F(X) \leq u) = P(X \leq F^{-1}(u))$$

$$= F(F^{-1}(u)) = u$$

FIGURA 17 – TEOREMA DO MÉTODO DA INVERSA DA PROBABILIDADE CUMULATIVA  
 FONTE: (DEVROYE, 1986, P. 28)

Gentle (2003, p. 103) ilustra a técnica por meio da Figura 18, onde se observa uma curva da função distribuição acumulada de uma função densidade qualquer, um valor  $u$  sorteado de uma distribuição uniforme de 0 a 1 e um valor  $x$  obtido da inversa da fda conhecida.

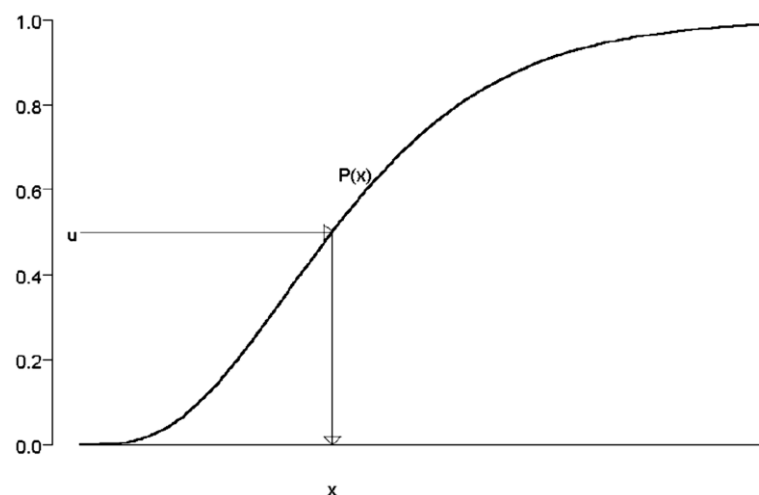


FIGURA 18 – EXEMPLO DE GRÁFICO DO MÉTODO DA INVERSA DA FDA  
 FONTE: (GENTLE, 2003, P. 103)

Outros métodos para transformação de dados de uma distribuição uniforme unitária para dados de outras distribuições podem ser encontrados em Devroye (1986) e Gentle (2003).

Para ilustrar o método de forma simples, pode-se tomar como exemplo a distribuição normal. Por meio do *software* R (R CORE TEAM, 2014) pode-se gerar um conjunto de números pseudoaleatórios (com distribuição normal) com base em uma semente. A Figura 19 ilustra a geração de 10 números pseudoaleatórios a partir da distribuição normal com média 10 e desvio padrão 5. Na linha 1 configura-se a semente do gerador *PRNG*. Na linha 2 solicita-se ao *software* que gere 10 números pseudoaleatórios da distribuição uniforme unitária, colocando o resultado na variável *n1*. Por padrão, o *software* utiliza o algoritmo “Mersenne-Twister” (MATSUMOTO e NISHIMURA, 1998) para geração desses números. Na linha 3 a variável *n1* é impressa apresentando os números gerados (linha 4 e 5). Na linha 6 é aplicado o método da inversa da fda para transformar os números pseudoaleatórios obtidos de uma distribuição uniforme para a distribuição normal. Nessa linha solicita-se ao *software* que retorne o valor da inversa da função cumulativa de probabilidade normal com média 10 e desvio padrão 0,5.

```
1 > set.seed(14)
2 > n1 <- runif(10)
3 > n1
4 [1] 0.2540337 0.6378273 0.9571886 0.5525467 0.9830671 0.5114739 0.9328234
5 [8] 0.4284277 0.4855851 0.3816522
6 > qnorm(n1,mean=10,sd=0.5)
7 [1] 9.669075 10.176329 10.859477 10.066049 11.060833 10.014382 10.748577
```

FIGURA 19 – COMANDO NO SOFTWARE R PARA GERAÇÃO DE NÚMEROS PSEUDOALEATÓRIOS DE UMA DISTRIBUIÇÃO NORMAL

No caso do exemplo ilustrado, a distribuição para a qual se deseja obter os números aleatórios é conhecida (distribuição normal).

Segundo Silverman (1986), a estimação de funções densidade de probabilidade pode ser feita por meio de métodos paramétricos ou não-paramétricos.

Os métodos paramétricos pressupõem que os dados a serem ajustados se encaixem numa distribuição conhecida. Por exemplo, pode-se estimar os valores da média ( $\mu$ ) e da variância ( $\sigma^2$ ) com base nas amostras e então substituir estas estimativas na equação da função de distribuição normal. (SILVERMAN, 1986). Além da função normal existem outras que podem ser ajustadas aos dados de uma amostra, tais como, Weibull 2P, Weibull 3P, log-normal, Gama, Beta, entre outras. Por outro lado, é necessário que se conheça a forma da distribuição dos dados para escolher as funções de distribuição que terão melhor aderência, sob pena de comprometer a qualidade do ajuste. Segundo Scott (2004), se as configurações experimentais forem bem conhecidas, então modelos paramétricos podem ser utilizados. Esse autor ainda cita como exemplo o caso de dados normais multivariados,  $N(\mu, \Sigma)$ , nos quais as características da densidade podem ser extraídas pela máxima verossimilhança dos parâmetros  $\mu$  e  $\Sigma$ . Entretanto, Scott (2004) alerta para o fato de que, se as configurações experimentais não forem bem conhecidas, ou se os dados não seguirem uma forma paramétrica conhecida, a estimação da função densidade por métodos não paramétricos deve ser aplicada. Portanto, se as amostras colocadas em um histograma não seguirem o formato de uma função de distribuição conhecida, não é recomendado aplicar métodos paramétricos para realizar o ajuste, mas devem-se aplicar métodos não paramétricos.

Algumas áreas do conhecimento têm utilizado o núcleo estimador de densidade, tais como mineração de dados, processamento de sinais e econometria. (SILVERMAN, 1986), (WAND e JONES, 1995), (JEFFERY, 1996) e (WOLFGANG *et al.*, 2004) (ALEXANDRE, 2009) citados por (GUIDOUM, 2013). Na área florestal pode-se encontrar alguns trabalhos que utilizaram a técnica do núcleo estimador de densidade para funções multivariadas, para identificação de risco de incêndio em florestas, como, por exemplo, o trabalho de Santos *et al* (2013).

### 2.3.2 Método do núcleo estimador variável

O núcleo estimador de densidade, ou KDE (Kernel Density Estimator) é uma técnica não paramétrica para a estimação da função densidade  $f(x)$  de uma variável aleatória  $X$ . Parzen (1962) propôs a formulação matemática para o núcleo estimador com base nos estudos anteriores de Fix e Hoges (1951) (publicado posteriormente em (FIX e HOGES, 1989)) e Rosenbatt (1956), conforme equação 3.

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - X_i}{h}\right) \quad (3)$$

Onde:

$\hat{f}(x)$  é a função que estima a função densidade  $f(x)$

$n$  é o número de unidades amostrais

$h$  é a janela ou parâmetro de suavidade

$K$  é uma função, denominada função núcleo

$x$  é o ponto onde é avaliada a função densidade

$X_i$  é a unidade amostral ou um dos valores da amostra aleatória  $X_1, X_2, \dots, X_n$  de uma distribuição univariada contínua

A função núcleo  $K$  é uma função densidade de probabilidade (fdp) com todas as propriedades de uma fdp e, por este motivo, a função  $\hat{f}(x)$  também será uma fdp. Para cada unidade amostral  $X_i$  uma função kernel é criada com o centro na amostra. Tal configuração garantirá que a função núcleo seja simétrica em relação à amostra. A fdp é então estimada pela soma de todas as funções núcleo e dividida pelo tamanho da amostra, assegurando, assim, as propriedades de uma fdp. Normalmente a função gaussiana é selecionada como função núcleo e, em casos assim, pode-se definir:

$$K\left(\frac{x - X_i}{h}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - X_i)^2}{2h^2}} \quad (4)$$

Substituindo (4) em (3) tem-se um estimador para a função densidade de probabilidade por meio do núcleo estimador com a função gaussiana como núcleo definida na equação 5.

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-X_i)^2}{2h^2}} \quad (5)$$

O parâmetro de suavidade  $h$ , também conhecido como janela, é determinante na qualidade do ajuste. Esse parâmetro pode ser comparado ao número de classes do histograma, mas não possui necessariamente o mesmo valor. Segundo Baranzano (2011), quando  $h$  é pequeno, resulta em uma estimativa superajustada; quando  $h$  é grande resulta em uma curva sobressuavizada. Scott (1992), citado por Sain (1996), afirma que o desempenho do estimador é dominado pela escolha do parâmetro de suavização. Sain (1996) afirma que a forma funcional da função núcleo é irrelevante se comparada à importância do parâmetro de suavidade.

Podem-se encontrar na literatura vários métodos para a otimização do parâmetro de suavidade  $h$ , entre eles, a regra de ouro de Silverman (1986), melhorada por Scott (1992), o método da validação cruzada tendenciosa e não tendenciosa de Scott e Terrel (1987) e o método de Sheather e Jones (1991).

Em especial, a regra de ouro de Silverman (1986) é definida na equação 6.

$$h = 0,9 \cdot \min \{ \sigma; IIQ \} \cdot n^{-\frac{1}{5}} \quad (6)$$

Onde:

$\sigma$  é o desvio padrão das unidades amostrais

$IIQ$  é o intervalo inter-quartil das unidades amostrais

$n$  é o número de unidades amostrais

Segundo Terrell e Scott (1992),  $h$  se mantém constante na maioria dos trabalhos aplicados, por outro lado pode-se variar  $h$  para melhorar a estimativa das densidades. Na literatura, esse método é conhecido como “núcleo estimador variável” ou “núcleo estimador adaptativo”.

Uma das implementações do núcleo estimador variável foi proposta por Silverman (1986), e é definida na equação 7.

$$\hat{f}_s(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h\lambda_i} K\left(\frac{x-X_i}{h\lambda_i}\right) \quad (7)$$



Onde:

$\hat{f}_s(x)$  é a função que estima a função densidade  $f(x)$

$\lambda_i$  é o fator de localidade (*"local bandwidth fator"*)

O fator de localidade tem a incumbência de tornar  $h\lambda_i$  pequeno nas regiões das modas e grandes nas caudas. Segundo Silverman (1986)  $\lambda_i$  é representado pela equação 8.

$$\lambda_i = \left( \frac{\tilde{f}(X_i)}{g} \right)^{-\alpha} \quad (8)$$

Onde:

$\tilde{f}(X_i)$  é uma estimativa piloto gerada pelo núcleo estimador fixo.

$\alpha$  é o parâmetro de sensibilidade,  $0 \leq \alpha \leq 1$

$g$  é a média geométrica de  $\tilde{f}(X_i)$

$$g = \left[ \prod_{i=1}^n \tilde{f}(X_i) \right]^{\frac{1}{n}} \quad (9)$$

O parâmetro  $\alpha$  controla a influência de  $\lambda_i$  na estimativa. Se  $\alpha = 0$ , então o método fica equivalente à abordagem do núcleo-estimador fixo. Pode-se afirmar, portanto, que o método núcleo estimador fixo é um caso particular do método núcleo estimador variável proposto por Silverman (1986) para  $\alpha = 0$ . O parâmetro  $h$  em (7) pode ser obtido da mesma maneira que o descrito em (6), ou utilizando outro método para a otimização do parâmetro de suavidade.

Segundo Nadaraya (1964), citado por Altman & Léger (1994), a função cumulativa de probabilidade do núcleo estimador pode ser definida pela equação 10.

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n L\left(\frac{x - X_i}{h}\right) \quad (10)$$

Aqui,  $L$  é uma função de distribuição acumulada de um núcleo positivo e  $L(x) = \int_{-\infty}^x K(t)dt$ .

Estendendo o conceito descrito por Nadaraya (1964) em (10), pode-se reescrever a função cumulativa de probabilidade do núcleo estimador variável, incluindo o fator de localidade  $\lambda_i$  conforme consta na equação 11.

$$\hat{F}(x) = \frac{1}{n} \sum_{i=1}^n L\left(\frac{x - X_i}{h\lambda_i}\right) \quad (11)$$

Para o caso em que o núcleo é uma função de distribuição gaussiana, conforme demonstrado na equação (12), pode-se reescrever a função (7) para (13):

$$K\left(\frac{x - X_i}{h\lambda_i}\right) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x - X_i)^2}{2h^2\lambda_i^2}} \quad (12)$$

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{\sqrt{2\pi}h\lambda_i} e^{-\frac{(x - X_i)^2}{2(h\lambda_i)^2}} \quad (13)$$

Isso a torna uma soma de funções de distribuição gaussianas com média  $X_i$  e desvio padrão  $h\lambda_i$ . Pode-se observar, que diferentemente do núcleo estimador fixo, neste caso o desvio padrão das funções gaussianas varia para cada amostra. Sabendo que a função de distribuição acumulada é definida pela integração da função de distribuição, tem-se que:

$$\begin{aligned} \hat{F}(x) &= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^x \frac{1}{\sqrt{2\pi}h\lambda_i} e^{-\frac{(x - X_i)^2}{2(h\lambda_i)^2}} \cdot \therefore \\ \hat{F}(x) &= \frac{1}{n} \sum_{i=1}^n F_{\mu=X_i; \sigma=h\lambda_i}(x) \end{aligned} \quad (14)$$

De forma análoga ao núcleo estimador fixo, a integração que está dentro do somatório em (14), nada mais é do que uma função de distribuição acumulada gaussiana com média em  $X_i$  e desvio padrão  $h\lambda_i$ . Portanto, neste caso, pode-se afirmar que a estimativa da distribuição acumulada é a soma das distribuições acumuladas de todas as curvas gaussianas definidas para cada uma das unidades amostrais  $X_i$ .

A inversa da probabilidade cumulativa pode ser obtida igualando-se o somatório das integrais à probabilidade conhecida. Por exemplo, para obter o valor de

$x$ , cuja probabilidade cumulativa é igual a 0,45 pode-se substituir  $\hat{F}(x) = P(X \leq x) = 0,45$  em (14) obtendo-se:

$$\begin{aligned}\hat{F}(x) &= 0,45 \therefore \\ 0,45 &= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^x \frac{1}{\sqrt{2\pi}h\lambda_i} e^{-\frac{(x-X_i)^2}{2(h\lambda_i)^2}} \therefore \\ 0,45 &= \frac{1}{n} \sum_{i=1}^n F_{\mu=X_i; \sigma=h\lambda_i}(x)\end{aligned}\tag{15}$$

Não é possível obter uma solução analítica para a integral de (14) ou (15), portanto é necessário recorrer a uma solução numérica a fim de se obter um valor de  $x$  para uma determinada probabilidade acumulada. Um dos possíveis algoritmos que pode ser aplicado para solucionar a equação (15) é o Brent (1973) presente no *software* R. O algoritmo *Generalized Reduced Gradient (GRC)* proposto por Lasdon, *et al.* (1978) e presente no suplemento Solver do *software* Microsoft Excel também pode ser aplicado para solucionar a equação (15).

No Apêndice 1 é apresentado o estudo dos momentos da função núcleo estimador variável para que seja possível o cálculo da estimativa da média e da variância para a função densidade de probabilidade.

No Apêndice 2 é apresentado um exemplo de aplicação do método núcleo estimador variável, com uma amostra de tamanho 10.

Dessa maneira, pode-se aplicar a inversa da probabilidade cumulativa com o método núcleo estimador variável a um conjunto de números pseudoaleatórios obtidos de uma distribuição uniforme, conforme descrito na seção 2.3.1, e obter-se um conjunto de números aleatórios da distribuição estimada por esse método.

### 3 MATERIAL E MÉTODOS

Este capítulo descreve a metodologia utilizada na construção e avaliação do *software*, as tecnologias empregadas no desenvolvimento do *software* e o método utilizado para geração dos dados pseudoaleatórios para cada um dos usuários (alunos) do *software*.

#### 3.1 METODOLOGIA PARA A CONSTRUÇÃO DO SOFTWARE EDUCACIONAL E BIBLIOTECA DE PROGRAMAÇÃO

Para o desenvolvimento do *software* educacional seguiu-se uma abordagem metodológica dupla: uma metodologia para a construção de material instrucional e técnicas de engenharia de *software* para a fase de desenvolvimento do *software*. Para a avaliação do *software* utilizou-se a pesquisa avaliativa por meio de um instrumento de avaliação com itens objetivos e questões abertas.

A metodologia para construção de material instrucional está baseada em um modelo de cinco fases, conhecido na literatura como modelo APDIA (Análise, Projeto, Desenvolvimento, Implementação e Avaliação - *ADDIE Model – Analysis, Design, Development, Implementation and Evaluation*). Forest (2014) afirma que esta metodologia foi concebida para o exército americano em 1975 pelo Centro de Tecnologia Educacional da Universidade Estadual da Flórida (*The Centre for Educational Technology Department at Florida State University*) e adaptada ao longo dos anos. Segundo Souza *et al* (2007), na fase de análise estudam-se as características do aprendiz e da tarefa a ser ensinada; a fase de projeto define a escolha da abordagem instrucional; o material é criado na fase de desenvolvimento; a implementação foca os aspectos de distribuição e a entrega do material; por último, na fase de avaliação verifica-se se o material atingiu os objetivos.

A construção do *software* educacional apresentado neste trabalho seguiu as seguintes etapas de desenvolvimento, com base no modelo APDIA. A Figura 20 ilustra cada uma das etapas do trabalho, segundo a metodologia definida nesta seção.

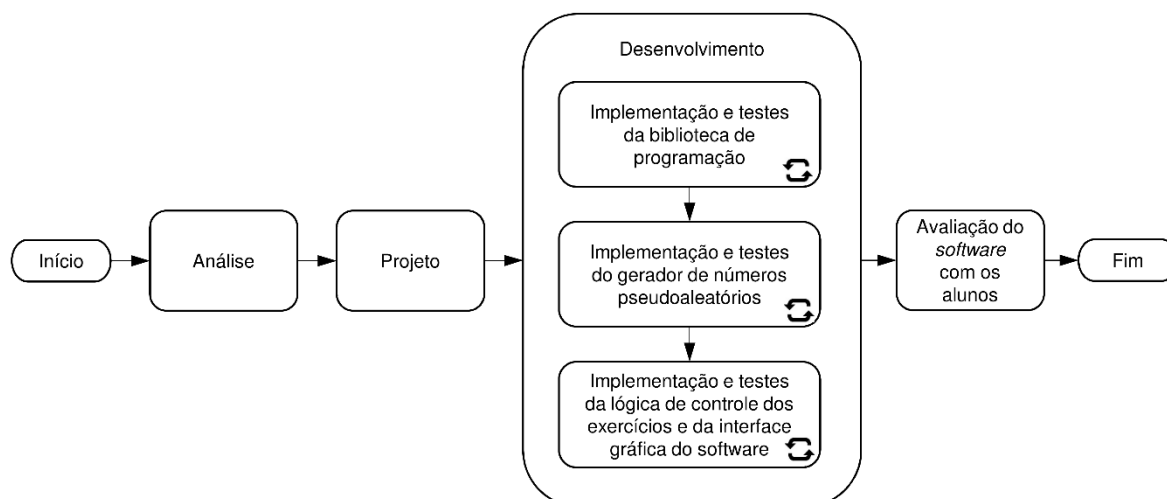


FIGURA 20 – FLUXOGRAMA DAS ETAPAS DO TRABALHO SEGUNDO A METODOLOGIA ADOTADA

### 3.1.1 Análise

Foi definido que o *software* abordaria a disciplina de inventário florestal e, mais especificamente exercícios práticos sobre os processos de amostragem revisados na seção 2.1 e, tendo como base teórica o livro texto Inventário Florestal (PÉLLICO NETTO e BRENA, 1997, p. 97-303). Nesta etapa definiu-se que as principais funcionalidades do *software*, listadas abaixo:

- Cada aluno deve ter o seu usuário e senha para utilização do sistema.
- Cada aluno deve ter o seu conjunto de dados para cada exercício do *software*.
- O *software* deve contemplar um exercício para cada um dos 16 processos de amostragem apresentados no livro texto Inventário Florestal (PÉLLICO NETTO e BRENA, 1997, p. 97-303).

- Cada exercício implementado terá um conjunto de questões, referentes aos cálculos estatísticos do processo de amostragem, que serão resolvidos pelo aluno.
- O software deve informar ao aluno se a resposta inserida para questão está correta ou incorreta.
- O *software* deve ter um sistema de pontuação por aluno, a fim de que cada aluno possa comparar o seu desempenho com os demais. O sistema de pontuação está descrito na seção 4.3.5.
- Deve ser possível exportar os dados utilizados na resolução dos exercícios em formato de planilha eletrônica.
- O software deve ser acessado de um navegador por meio de uma conexão com a Internet.

### 3.1.2 Projeto

Nesta fase as tecnologias a serem utilizadas no *software* foram definidas (apresentadas na seção 3.2). Concernente à tecnologia foram realizados testes para definir a utilização das bibliotecas mais apropriadas para o *software*. Os modelos de classes que suportaram o desenvolvimento da biblioteca de programação para processos de amostragem de inventário florestal foram iniciados nessa etapa e seguiram todos os conceitos revisados na seção 2.2.2. A modelagem das classes e a aplicação dos conceitos serão apresentados na seção 4.1.

Nessa etapa foi definido que seria necessário utilizar algum método para geração aleatória de dados para que cada aluno tivesse o seu próprio conjunto de dados. A biblioteca de programação Apache Commons Math 3.3.0 (THE APACHE SOFTWARE FOUNDATION, 2014) que, segundo a sua documentação, implementa o método núcleo estimador variável, foi utilizada para geração pseudoaleatória dos

dados deste trabalho. A utilização da biblioteca de programação e o método para geração aleatória de dados está descrito na seção 3.3.

### 3.1.3 Desenvolvimento

Nesta fase utilizou-se a técnica, conhecida na área de engenharia de *software* como prototipação ou prototipagem (ver seção 2.2.1). Conforme demonstrado na Figura 20 essa fase foi dividida em 3 atividades distintas que foram realizadas em iterações, ou seja, foram desenvolvidas e testadas em ciclos, inclusive com validações com os professores orientadores do trabalho. Pode-se observar, na Figura 20, que em cada atividade existe um símbolo no canto inferior direito representando a iteratividade dessas tarefas. As três atividades dessa fase foram definidas como:

- Implementação e testes da biblioteca de programação: O primeiro protótipo gerado foi uma biblioteca orientada a objetos para cálculos estatísticos dos processos de amostragem florestal. Os modelos de classes foram aprimorados e estão apresentados na seção 4.1. A validação dessa biblioteca se deu com a resolução automática dos exercícios do livro texto (PÉLLICO NETTO e BRENA, 1997, p. 97-303) e com a execução de testes unitários. As evidências dos testes realizados são apresentadas no Apêndice 5. Pode-se afirmar que uma das orientações dos métodos ágeis (revisados na seção 2.2.1) diz respeito aos testes que devem ser executados de maneira contínua e automática, o que se adotou na construção da biblioteca de programação para inventário florestal.
- Implementação e testes do gerador de números pseudoaleatórios: Os protótipos gerados nessa etapa tinham o objetivo de gerar dados com base nos conjuntos de dados do livro texto utilizado (PÉLLICO NETTO e BRENA, 1997, p. 97-303). A implementação desses protótipos foi feita com base no método para geração aleatória de dados descrito na seção 3.3. Foi

realizado o teste de Kolmogorov-Smirnov para verificar se as unidades amostrais geradas e as unidades amostrais da população que deram origem aos dados podem ser consideradas da mesma distribuição probabilística. O método de teste está descrito na seção 3.3.

- Implementação e testes da lógica de controle dos exercícios e da interface gráfica do software: Os protótipos seguintes tiveram a intenção de validar as interações com os usuários (alunos). Portanto essa atividade de construção do *software* se deu por meio de iterações, em que o objetivo de cada iteração era validar o protótipo com um autor do livro texto e o orientador desta pesquisa. O funcionamento das telas do *software* está apresentado na seção 4.3 deste documento.

#### 3.1.4 Implantação

Depois de construído o *software* foi implantado em uma máquina servidora na Universidade Federal do Paraná. Neste caso a distribuição do *software* é feita de forma *on-line*, uma vez que não é necessária a instalação.

#### 3.1.5 Avaliação

Nesta fase foi realizada uma pesquisa avaliativa baseada em um instrumento de avaliação aplicado aos alunos do 5º período do curso de Engenharia Florestal da UFPR no primeiro semestre de 2014. No primeiro encontro essa turma foi exposta ao *software* por meio de uma demonstração. No encontro seguinte, pediu-se aos alunos que resolvessem o desafio 6 do *software* (ver seção 4.3) em um laboratório com máquinas individuais. Por último foi solicitado à turma responder ao instrumento de avaliação. A prática de avaliar um objeto de aprendizagem por meio da aplicação de instrumentos de avaliação não é incomum, podendo ser observada em alguns



trabalhos científicos (KAY e KNAACK, 2007). A metodologia para a avaliação do software está detalhada na seção 3.3.

O Quadro 22 apresenta a estrutura do instrumento de avaliação que foi aplicado. Esse instrumento de avaliação é uma adaptação do que foi aplicado nas pesquisas de Kay e Knaack (2007) e Nugent, Soh e Samal (2006). Os primeiros fizeram uma avaliação de 5 objetos de aprendizagem com 30 professores e 221 estudantes de 12 escolas secundárias diferentes. Nugent, Soh e Samal (2006) construíram e fizeram uma avaliação de um objeto de aprendizagem para alunos do ensino superior em Ciência da Computação.

No.	Tipo	Itens
1	Objetiva	O Simulador Florestal agrega valor à disciplina de Inventário Florestal.
2	Objetiva	O Simulador Florestal possibilita outra estratégia de aprendizado.
3	Objetiva	Eu sinto que o Simulador Florestal pode beneficiar o entendimento sobre os cálculos estatísticos referentes aos processos de amostragem.
4	Objetiva	O <i>software</i> Simulador Florestal é de fácil operação.
5	Objetiva	Outros exercícios ao longo do curso poderiam ser apresentados por um objeto de aprendizagem semelhante a este.
6	Objetiva	De um modo geral o Simulador Florestal é um bom complemento ao conteúdo ministrado em sala de aula.
7	Objetiva	O sistema de pontuação do <i>software</i> me incentivou a pensar melhor na solução dos problemas propostos.
8	Objetiva	A correção dos exercícios em tempo real me ajudou na solução dos problemas.
9	Objetiva	Eu estou interessado em utilizar o Simulador Florestal novamente.
10	Objetiva	De um modo geral, como você avalia o Simulador Florestal. (Péssimo, Fraco, Regular, Bom e Excelente)
11	Aberta	O que você gostou no Simulador Florestal?
12	Aberta	O que você não gostou no Simulador Florestal?
13	Aberta	Você tem alguma sugestão sobre o Simulador Florestal?

QUADRO 22 – INSTRUMENTO DE AVALIAÇÃO  
FONTE: O AUTOR (2014)

Nos itens objetivos, cada estudante marcou uma das seguintes opções: discordo, discordo parcialmente, neutro, concordo parcialmente e concordo.

Os resultados dos itens objetivos foram avaliados e discutidos segundo o resultado da média e desvio padrão para cada um. Foi atribuído um peso para cada opção: 1 para discordo, 2 para discordo parcialmente, 3 para neutro, 4 para concordo parcialmente e 5 para concordo. Com base nesses pesos por item, foram calculadas a média e o desvio padrão para verificar o nível de concordância média de cada item

e também o nível de dispersão que cada um gerou nas respostas. Os itens abertos foram utilizadas para uma avaliação qualitativa do objeto de aprendizagem.

### 3.2 TECNOLOGIAS UTILIZADAS

Para a programação do núcleo do *software* foi utilizada a linguagem Java versão 7, que é uma linguagem de propósito geral, a qual segue o paradigma orientado a objetos e possui uma biblioteca de programação bastante ampla. A Oracle (2014), empresa que mantém a linguagem atualmente, define Java como uma linguagem de programação e uma plataforma de desenvolvimento. Essa linguagem foi escolhida por ser amplamente utilizada no mercado e também na academia. Segundo o índice Tiobe (TIOBE SOFTWARE, 2014), que mede a popularidade das linguagens de programação, Java é a segunda em popularidade no mês de março de 2014, com 16,4%. Entretanto, nos últimos 12 anos essa linguagem está alternando entre a primeira e segunda posição em popularidade. Outra justificativa para a escolha da linguagem é o seu suporte nativo para a construção de aplicações para Internet.

No desenvolvimento da interface gráfica, foi utilizada a linguagem JavaScript, que é interpretada pelo navegador da Internet. A utilização dessa linguagem em conjunto com a biblioteca ExtJs versão 3.4 (SENCHA, 2014) possibilitou o desenvolvimento de uma interface gráfica interativa e disponibilizada via navegador da Internet. Essa linguagem e a biblioteca foram escolhidas porque os navegadores atuais suportam de forma nativa a linguagem JavaScript, sem a necessidade de instalação de *softwares* adicionais. Além disso, foi possível acelerar o desenvolvimento do *software*.

Para a codificação do software foi utilizado o “*Eclipse Java EE IDE for Web Developers*” que é um ambiente de programação integrado (*IDE – Integrated Development Enviroment*) voltado ao desenvolvimento de aplicações corporativas (EE

– Enterprise Edition) para a internet na linguagem Java. Foi utilizada a versão “Kepler Service Release 1” desse ambiente.

Em complemento às linguagens de programação, foram utilizadas algumas bibliotecas que facilitaram e aceleraram o desenvolvimento do núcleo do *software*, tais como: Apache Commons Math 3.3.0 (biblioteca que contém componentes matemáticos e estatísticos, inclusive o método núcleo estimador variável para geração de dados pseudoaleatórios), Apache Commons Bean Utils 1.8.0, Apache Commons Codec 1.5, Apache Commons Collections 3.2.1, Apache Commons Lang 2.5, Apache Commons Logging 1.1.1, EzMorph 1.0.6 e Apache Poi 3.9 (THE APACHE SOFTWARE FOUNDATION, 2013), (EZMORTH, 2014), (THE APACHE SOFTWARE FOUNDATION, 2014).

O armazenamento de dados foi suportado pelo *software* MySQL versão 5.6 (ORACLE, 2014), que é um sistema gerenciador de banco de dados relacional e que utiliza a linguagem SQL (*Structured Query Language* – Linguagem de Consulta Estruturada) para manipulação dos dados.

O *software* desenvolvido está disponível na Internet, e foi necessário utilizar o Tomcat (versão 7) (THE APACHE SOFTWARE FOUNDATION, 2014), que é um servidor para páginas dinâmicas e suporta a linguagem Java.

O diagrama da Figura 21 ilustra os componentes de tecnologia utilizados no *software* e seus relacionamentos com o *hardware*. O navegador, a biblioteca ExtJs e a linguagem JavaScript executam na máquina cliente (máquina do aluno). O banco de dados (MySQL), o servidor de internet (Apache Tomcat), as bibliotecas (Apache Commons) e os componentes escritos na linguagem Java executam na máquina servidora.

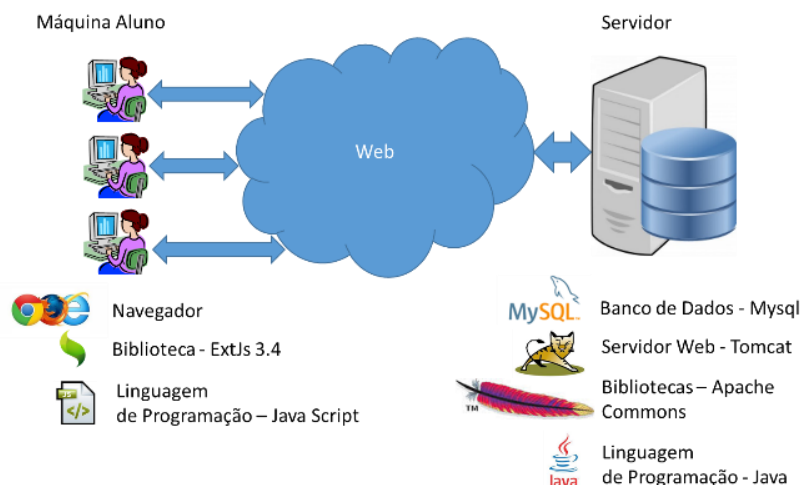


FIGURA 21 – DIAGRAMA SIMPLIFICADO DE COMPONENTES DO SOFTWARE  
 FONTE: O AUTOR (2014)

### 3.3 MÉTODO PARA GERAÇÃO ALEATÓRIA DE DADOS

Segundo Deshpande e Huang (2009), autores de livros e desenvolvedores de jogos poderiam trabalhar juntos para disponibilizar materiais apropriados como suplementos dos livros texto. Expandindo esse conceito, autores e desenvolvedores de *software* poderiam trabalhar de maneira integrada para produzir objetos de aprendizagem como material complementar aos livros texto.

O objeto de aprendizagem desenvolvido no presente trabalho teve, como base teórica, os 16 processos de amostragem apresentados no livro texto Inventário Florestal (PÉLLICO NETTO e BRENA, 1997, p. 97-303) e revisados na seção 2.1, cujos dados utilizados constam da obra citada.

Por outro lado, os dados serviram como base para que o sistema gerasse automaticamente subconjuntos aleatórios individuais por aluno. Para isso foi utilizada a biblioteca de programação Apache Commons Math 3.3.0 (THE APACHE SOFTWARE FOUNDATION, 2013). Essa biblioteca estima uma função densidade de probabilidade por meio do método não paramétrico “núcleo estimador variável” e gera dados pseudoaleatórios para serem utilizados como volumes por hectare de unidades

amostrais de área fixa. A geração de dados pseudoaleatórios e o método do núcleo estimador variável foram revisados na seção 2.3.

Para o entendimento e demonstração do método de geração de dados pseudoaleatórios utilizando o “núcleo estimador variável” foi realizado um experimento no software R com um número limitado de unidades amostrais (10 unidades), e implementada algumas funções nesse software. Esse experimento, descrito no Apêndice 2, tem a intenção de demonstrar o funcionamento do método, mas as funções implementadas para ele não foram utilizadas no *software* educacional desenvolvido. O software desenvolvido utilizou a biblioteca de programação na linguagem Java chamada Apache Commons Math 3.3.0 (THE APACHE SOFTWARE FOUNDATION, 2013), pois esta, segundo a sua documentação, já implementa o método núcleo estimador variável para geração de números pseudoaleatórios.

A fim de demonstrar o procedimento para geração de dados pseudoaleatórios no software educacional, a seguir estão apresentados os passos que foram seguidos para gerar os dados utilizados nos primeiros 9 exercícios desenvolvidos no software. O procedimento de geração de dados pseudoaleatórios para os outros 7 exercícios foi o mesmo, variando os dados de origem e algumas propriedades da estrutura da população. O Quadro 23 demonstra para cada exercício desenvolvido, quais foram os dados de origem utilizados do livro texto (PÉLLICO NETTO e BRENA, 1997, p. 97-303), com o número do quadro e a página.

<b>Exercício</b>	<b>Quadro e Página</b>
1, 2, 3, 4, 5, 6, 7,8 e 9	Quadro 4.1, página 98
10	Quadro 8.3, página 197
11	Quadro 9.1, página 222
12	Quadro 10.1, página 232
13	Quadro 11.1, página 245
14	Quadro 11.2, página 255
15	Quadro 11.3, página 273
16	Quadro 11.6, página 293

QUADRO 23 – DADOS DE ORIGEM UTILIZADOS DO LIVRO TEXTO (PÉLLICO NETTO E BRENA, 1997, P. 97-303) POR EXERCÍCIO IMPLEMENTADO  
FONTE: O AUTOR (2104)

A Figura 22 ilustra os passos implementado no *software* educacional para geração de dados pseudoaleatórios.

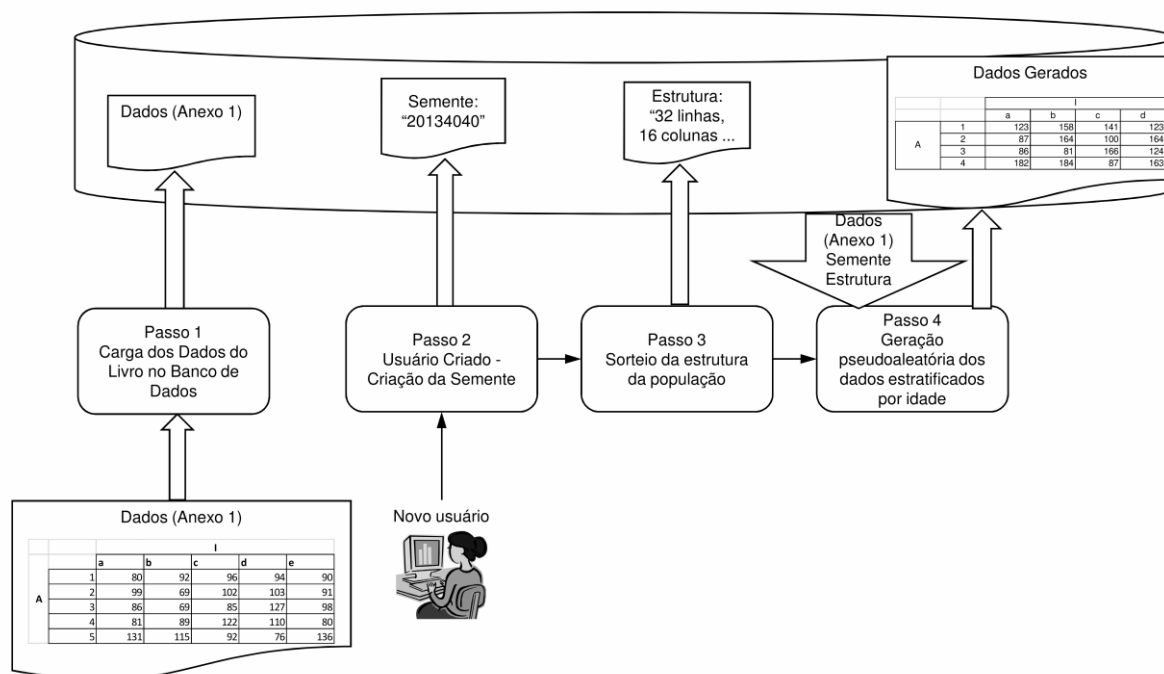


FIGURA 22 – PASSOS IMPLEMENTADOS NO SOFTWARE EDUCACIONAL PARA GERAÇÃO DE DADOS PSEUDOALEATÓRIOS  
FONTE: O AUTRO (2104)

Cada um dos passos apresentados na Figura 22 estão descritos a seguir.

### Passo 1 – Carga dos dados

Foi realizada uma carga dos dados apresentados no Anexo 1 para o banco de dados. Esses dados foram carregados uma única vez na tabela “unidades amostrais de área fixa” (ver modelo E-R no ).

### Passo 2 – Criação da semente

Para cada usuário (aluno) criado no sistema é obtido o número da matrícula do aluno como semente (ver definição de semente na seção 2.3). Na UFPR o número de matrícula dos alunos é formado por uma sequência de 3 letras e 8 números. Para formar a semente com um número inteiro retirou-se as 3 letras da matrícula, ficando apenas com os 8 números gerados. Por exemplo, para um aluno com matrícula GRR20143030, a semente seria 20143030.

### Passo 3 – Sorteio da estrutura da população

Além dos dados propriamente dito, algumas propriedades da estrutura da população também são sorteadas com base na distribuição uniforme. A estrutura da população tomada como base para os exercícios 1 a 7 está descrita no Anexo 1. O Quadro 24 ilustra quais propriedades da estrutura dessa população são sorteadas e quais são os possíveis valores (domínio) assumidos por cada propriedade.

As propriedades “quantidade de linhas primárias” e “quantidade de colunas primárias” são definidas em função do sorteio da “quantidade de linhas” e da “quantidade de colunas”, pois devem ser múltiplos.

Propriedade	Domínio
Quantidade de linhas secundárias	30, 32 ou 35
Quantidade de colunas secundárias	15, 16 ou 18
Comprimento da unidade amostral	25 ou 50 metros
Largura da unidade amostral	20 ou 30 metros
Quantidade de linhas primárias	6 para 30 linhas 7 para 35 linhas 8 para 32 linhas
Quantidade de colunas primárias	3 para 15 e 18 linhas 4 para 16 linhas

QUADRO 24 – PROPRIEDADES DA ESTRUTURA DA POPULAÇÃO QUE SÃO SORTEADOS  
FONTE: O AUTOR (2014)

### Passo 4 – Geração pseudoaleatória dos dados estratificados por idade

O total de valores pseudoaleatórios que são gerados para cada aluno é calculado pelo produto entre “quantidade de linhas secundárias” e “quantidade de colunas secundárias”.

Para que os dados gerados ficassem com valores mais próximos dos dados de origem, optou-se por fazer a geração de dados separada por idade. No Anexo 1 tem-se a identificação da idade de cada unidade amostral pela cor da célula. Portanto, nesse caso, gera-se três conjuntos de dados, um para cada idade (6, 9 e 12 anos). O sistema também sorteia a quantidade de dados geradas para as idades 6 e 9 anos, dentro de um limite proporcional que foi arbitrado entre 29 e 36% do total de valores pseudoaleatórios. A quantidade de dados geradas para 12 anos será obtida em função das outras duas e completa os 100% do total de valores.

Por último, o sistema, utilizando a classe “*EmpiricalDistribution*” da biblioteca de programação da Apache (THE APACHE SOFTWARE FOUNDATION, 2013), como os dados carregados do Anexo 1 (passo 1), com a semente gerada por aluno (passo 2), gera e armazena uma nova população de dados pseudoaleatórios para cada idade, respeitando a estrutura definida no passo 3. Essa nova população de dados foi gerada a partir do método núcleo estimador variável, considerando cada uma das três (6, 9 e 12 anos) distribuições formada pelos dados originais. A classe “*EmpiricalDistribution*” implementa internamente todos os passos necessários para geração de dados pseudoaleatórios por meio do método núcleo estimador variável. Com o objetivo de ilustrar o funcionamento, o método núcleo estimador variável foi estudado neste trabalho (seção 2.3.2) e um exemplo com um número de dados limitados foi apresentada no Apêndice 2 utilizando o software R.

Os dados gerados, bem como os dados referentes às informações do desempenho de cada aluno ao resolver os exercícios propostos pelo *software*, são armazenados no banco de dados relacional MySQL versão 5.6 (ORACLE, 2014) instalado na máquina servidora (Figura 21). O modelo físico do banco de dados está ilustrado no Apêndice 3 deste trabalho.

### Teste de Kolmogorov-Smirnov

Com o objetivo de testar se os dados gerados com o método descrito nessa seção respeitam a distribuição dos dados de origem foi realizado o teste de Kolmogorov-Smirnov.

O teste de Kolmogorov-Smirnov, consiste em comparar a máxima diferença entre a frequência relativa acumulada de uma amostra, e a frequência relativa acumulada de outra amostra, conforme demonstrado em (16).

$$d_{cal} = \max |F(X) - G(X)| \quad (16)$$

Em que  $F(X)$  é a função de distribuição acumulada empírica dos dados de origem e  $G(X)$  é a função de distribuição acumulada empírica dos dados gerados pelo método do núcleo estimador variável.



A hipótese da nulidade é aceita se ambas as amostras, no caso desse trabalho as amostras de origem e gerada, são provenientes de uma mesma distribuição. Rejeita-se a hipótese da nulidade (à um nível de significância  $\alpha$ ) se o valor calculado for maior do que valor tabelado (ou crítico), portanto se  $d_{cal} > d_{tab}$ .

O valor ( $d_{tab}$ ) crítico calculado para o nível de significância  $\alpha=0,05$  é obtido pela equação apresentada em (17).

$$d_{tab} = 1,36 \cdot \sqrt{\frac{m+n}{m \cdot n}} \quad (17)$$

Em que  $m$  é o número de unidades amostrais presentes nos dados de origem e  $n$  é o número de unidades amostrais que foram gerados pelo método núcleo estimador variável.

Para se gerar as funções de distribuição acumulada empírica  $F(X)$  e  $G(X)$ , traça-se o histograma dos dados de origem, obtém-se o tamanho das classes, as frequências correspondentes por classes e acumula-se a frequência relativa. Para realizar essa tarefa utilizou-se a função `hist` do software R (R CORE TEAM, 2014), que calcula o número de classes e os tamanhos das classes baseado no método de Sturges (1926), definido por:  $nc = 1 + 3,3 \cdot \log(n)$ .

## 4 RESULTADOS E DISCUSSÃO

Este capítulo apresenta os principais resultados obtidos neste trabalho. A seção 4.1 apresenta a modelagem das classes e seus relacionamentos da biblioteca de programação para os processos de amostragens, a seção 4.2 apresenta o gerador de dados pseudoaleatórios, a seção 4.3 descreve as funcionalidades do sistema desenvolvido e a seção 4.4 apresenta o resultado da avaliação realizada com os alunos do *software* desenvolvido.

### 4.1 BIBLIOTECA DE PROGRAMAÇÃO PARA PROCESSOS DE AMOSTRAGEM EM INVENTÁRIO FLORESTAL

A implementação de uma biblioteca de programação para processos de amostragem tornou-se necessária antes do desenvolvimento do *software* propriamente dito. Para que isso fosse viável, elaborou-se um modelo orientado a objetos (ver seção 2.2.2) para suportar os cálculos estatísticos envolvido nos processos de amostragem (ver seção 2.1).

A biblioteca desenvolvida possui 31 classes, sendo que as principais serão apresentadas e comentadas nessa seção por meio de diagramas de classes. Todas as classes que foram desenvolvidas para a biblioteca de programação estão apresentadas no Apêndice 4.

Para simplificar os diagramas, optou-se por apresentar os atributos e métodos de cada classe somente na primeira vez em que essa classe aparece em um diagrama. São apresentados somente os métodos públicos de cada classe. Por motivo de organização, optou-se por apresentar vários diagramas no lugar de um grande diagrama.

#### 4.1.1 Modelagem da unidade amostral

A primeira etapa consistiu em representar os processos de amostragem em um modelo computacional orientado a objetos e em identificar as entidades que fazem parte desses processos. Neste ponto, o conceito de Unidade Amostral (ver seção 2.1) revelou-se como a principal entidade a ser manipulada pelo sistema, uma vez que todos os cálculos tomam por base um conjunto de unidades amostrais. Portanto, utilizando o conceito de classe abstrata (ver seção 2.2.2) modelou-se uma Unidade Amostral por meio da classe *UnidadeAmostral*, conforme ilustrado na Figura 23.

Essa modelagem garante que a classe Unidade Amostral possa ser utilizada pelo método de área fixa, método de Bitterlich, método de Strand, método de Prodan, ou qualquer outro método em que os seguintes estimadores sejam definidos: número de árvores, número de árvores por hectare, área basal, área basal por hectare, volume e volume por hectare. Isso é possível, pois os métodos da classe que retornam os estimadores (`getNumArvores()`, `getNumArvoresPorHectare()`, `getAreaBasal()`, `getAreaBasalPorHectare()`, `getVolume()` e `getVolumePorHectare()`) são abstratos, portanto não possuem implementação. Essa ideia está de acordo com Péllico e Brena (1997, p. 18), os quais afirmam que as unidades amostrais, em inventários florestais, podem ser constituídas por parcelas de área fixa, pontos amostrais ou árvores.

<b><i>UnidadeAmostral</i></b>
<code>+UnidadeAmostral()</code> <code>+getIdBD():long</code> <code>+getId():String</code> <code>+getNumArvores():double</code> <code>+getNumArvoresPorHectare():double</code> <code>+getAreaBasal():double</code> <code>+getAreaBasalPorHectare():double</code> <code>+getVolume():double</code> <code>+getVolumePorHectare():double</code> <code>+getValor(String):double</code>

FIGURA 23 – CLASSE QUE REPRESENTA A UNIDADE AMOSTRAL  
 FONTE: O AUTOR (2014)

#### 4.1.2 Modelagem da unidade amostral de área fixa

No caso do sistema desenvolvido e avaliado, utilizou-se somente o método de área fixa; por esse motivo, implementou-se somente esse tipo de Unidade Amostral, que é chamada no modelo por Unidade Amostral de Área Fixa (Figura 24).

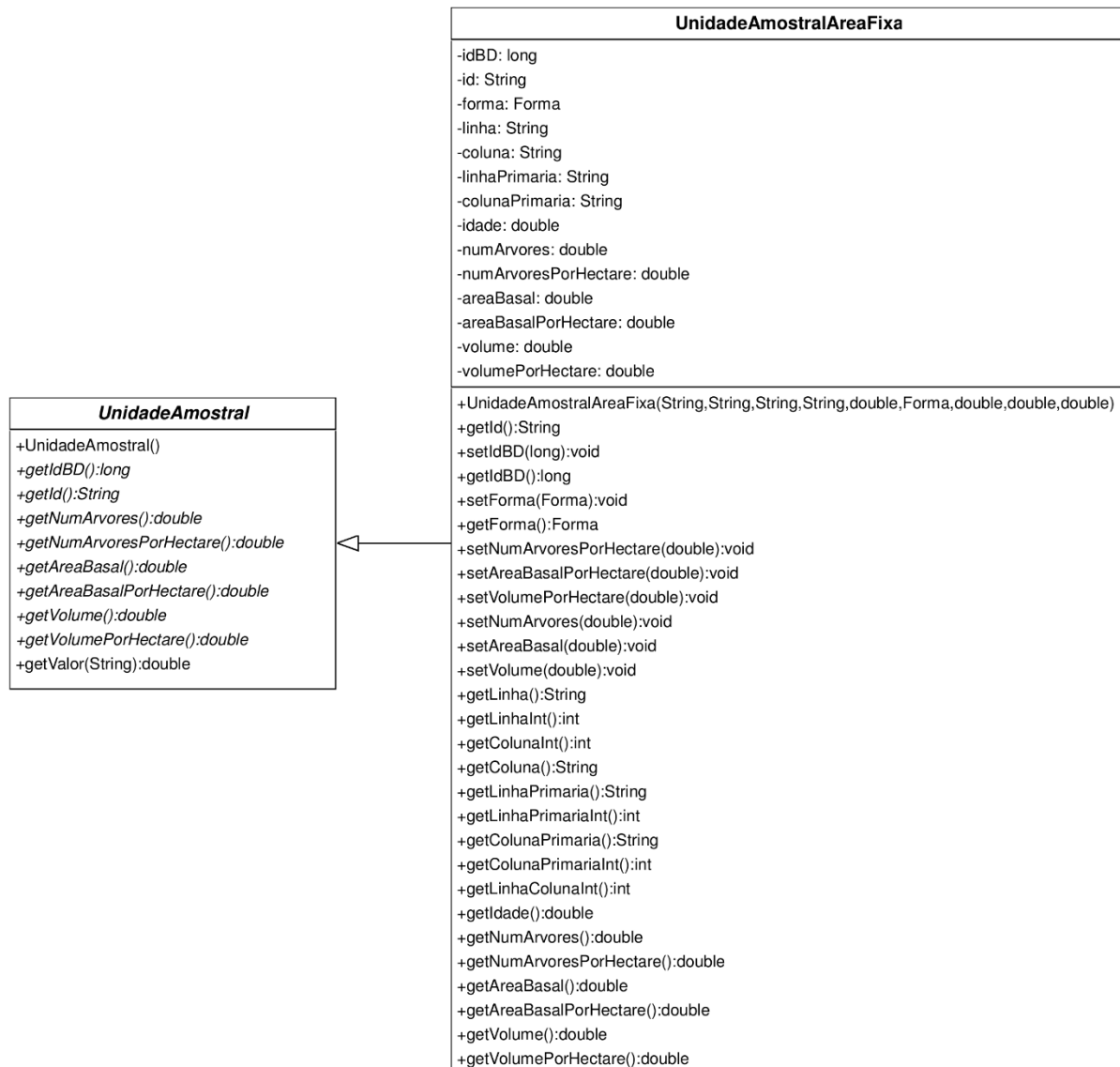


FIGURA 24 – RELAÇÃO ENTRE AS CLASSES QUE REPRESENTAM A UNIDADE AMOSTRAL E A UNIDADE AMOSTRAL DE ÁREA FIXA  
FONTE: O AUTOR (2014)

Pode-se observar na Figura 24 uma relação de herança (ver seção 2.2) entre a classe *UnidadeAmostrai* e *UnidadeAmostraiAreaFixa* (Figura 24). No diagrama da

Figura 24 a seta que aponta para a classe *UnidadeAmostrai* indica que *UnidadeAmostraiAreaFixa* é uma especialização da classe *UnidadeAmostrai*. Nesse caso, a classe *UnidadeAmostrai*, é abstrata, portanto, a classe *UnidadeAmostraiAreaFixa* deve fornecer uma implementação para todos os métodos abstratos da classe herdada. A classe *UnidadeAmostraiAreaFixa* não possui nenhum método abstrato por se tratar de uma classe real.

Apesar de o sistema tratar somente do método de área fixa, por meio da implementação de outra classe, é possível incluir outros métodos de amostragem. Para isso, basta que um programador crie uma nova classe que tenha uma relação de herança com a classe *UnidadeAmostrai*. Por exemplo, se um programador desejar implementar o método de Bitterlich, deverá criar uma nova classe chamada *UnidadeAmostraiBitterlich* que tenha uma relação de herança com a classe *UnidadeAmostrai* e implementar os estimadores para cada um dos métodos abstratos.

#### 4.1.3 Modelagem da amostra

Outro importante conceito modelado nessa biblioteca de programação refere-se ao conjunto de unidades amostrais, denominado de Composição de Unidades Amostrais (*ComposicaoUnidadesAmostrais*). Pode-se comparar a modelagem a uma amostra que, segundo Péllico e Brena (1997, p. 16), pode ser definida como uma parte da população, constituída de indivíduos que apresentam características comuns, as quais identificam a população a que pertencem.

Por outro lado, no caso das amostragens estratificadas, este mesmo conjunto de unidades amostrais pode se comportar como uma unidade amostral, portanto a entidade Composição de Unidades Amostrais possui um conjunto de unidades amostrais, mas também pode definir os estimadores: número de árvores, número de árvores por hectare, área basal, área basal por hectare, volume e volume por hectare.

Na Figura 25 podem-se observar as duas relações descritas entre as classes *ComposicaoUnidadesAmostrais* e *UnidadeAmostrai*. Por meio de uma flecha vazada está caracterizada a relação de associação entre as duas classes, na qual tem-se que a classe *ComposicaoUnidadesAmostrais* possui um conjunto de unidades amostrais. A outra relação (flecha fechada) indica uma relação de herança entre as duas classes, portanto, podemos afirmar que a classe *ComposicaoUnidadesAmostrais* é também uma *UnidadeAmostrai*.

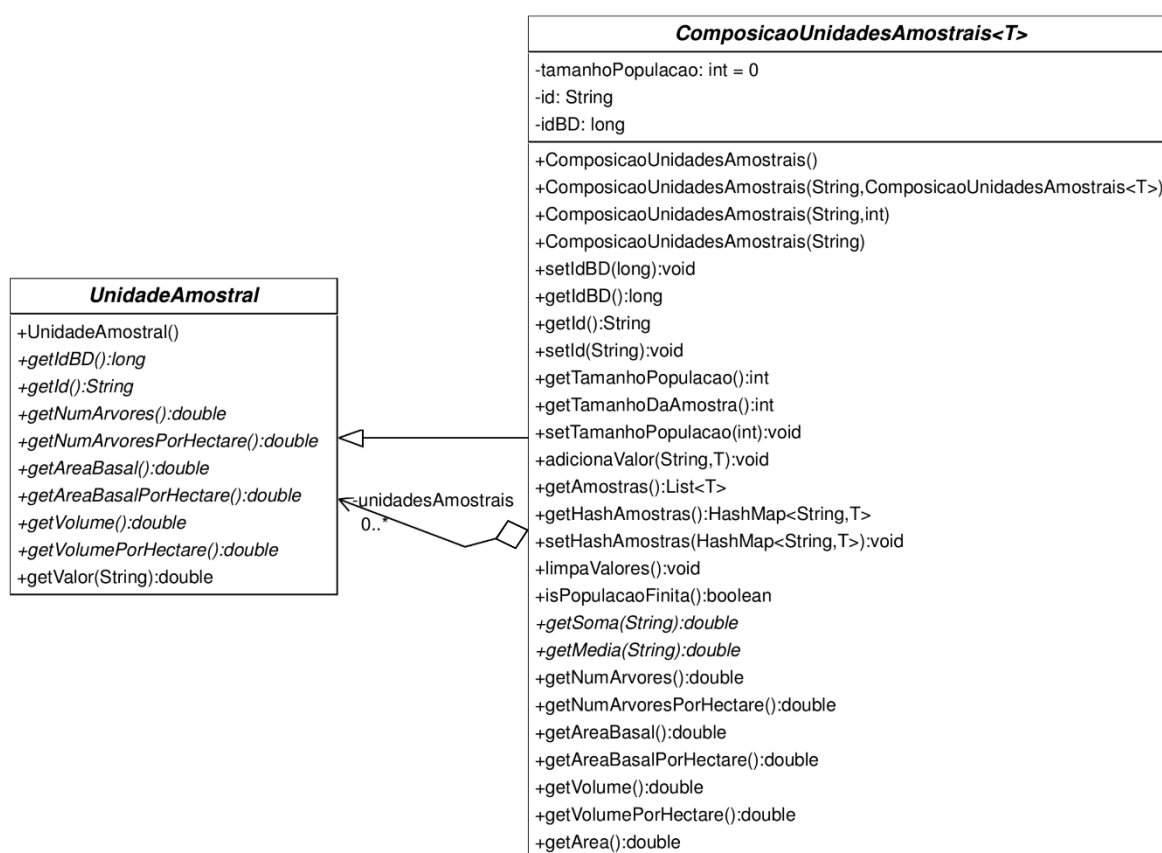


FIGURA 25 – RELAÇÃO ENTRE AS CLASSES QUE REPRESENTAM A AMOSTRA E SUAS UNIDADES AMOSTRAIS  
FONTE: O AUTOR (2014)

Pode-se observar que a classe *ComposicaoUnidadesAmostrais* é abstrata, com dois métodos abstratos: *getSoma* e *getMedia*. Os métodos retornam respectivamente a soma e a média da variável requerida. Mas a implementação desses métodos dependerá do tipo de processo de amostragem envolvido, por isso são métodos abstratos.

Outra característica importante dessa classe é o fato de utilizar o conceito de classe genérica. No diagrama da Figura 25, pode-se perceber essa característica por meio da letra T escrita entre os símbolos menor (<) e maior (>). Por meio desse conceito é possível definir composições de unidades amostrais de qualquer tipo, ou seja, essa classe pode representar um conjunto de unidades amostrais por parcelas de área fixa ou de pontos amostrais.

A classe foi modelada dessa maneira para que os processos de amostragem pudessem ser definidos a partir dela.

A classe *ComposicaoUnidadesAmostraisVO* foi criada como uma especialização da classe *ComposicaoUnidadesAmostrais* parametrizando o tipo genérico <T> para o tipo *UnidadeAmostralAreaFixa*. A classe *ComposicaoUnidadesAmostraisVO* é utilizada para criar uma amostra simples de unidades amostrais que pudesse ser utilizada por todas as classes dos processos de amostragem.

#### 4.1.4 Modelagem dos processos de amostragem

A seguir estão descritas as modelagens das entidades dos processos de amostragem implementados nessa biblioteca de programação.

##### **Amostragem simples**

A classe *AmostragemSimple* (Figura 26) foi definida como uma especialização da classe *ComposicaoUnidadesAmostrais*, sendo que o tipo de unidade amostral é o de área fixa. Então, pode-se afirmar que a classe *AmostragemSimple* é uma composição de unidades amostrais do tipo área fixa. Pode-se observar esta relação entre as classes na Figura 26, que não representa todos os métodos e atributos das classes por uma questão de apresentação.

Essa classe é utilizada para os cálculos das estimativas de um conjunto de unidades amostrais da população que tenham igual chance de participar da amostra,

caracterizando, assim, o processo de amostragem aleatória simples. Entre as estimativas calculadas por essa classe estão: média, desvio padrão, coeficiente de variação, erro padrão, fração de amostragem, variância da média, variância da média relativa, erro absoluto, erro relativo, intervalo de confiança para a média e para o total, total estimado e intensidade de amostragem.

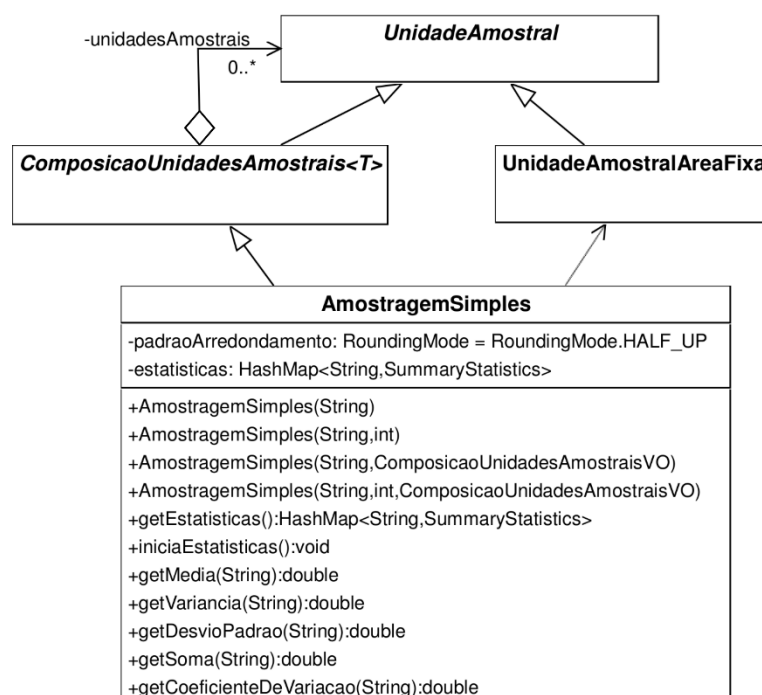


FIGURA 26 – CLASSE QUE REPRESENTA A AMOSTRAGEM SIMPLES E SUAS RELAÇÕES  
FONTE: O AUTOR (2014)

## Parâmetros da População

A classe que calcula os parâmetros de uma população é uma especialização da classe *AmostragemSimples*, pois os métodos utilizados para calcular as estimativas de uma amostra têm a mesma implementação dos métodos utilizados para calcular os parâmetros de uma população. Pode-se observar essa relação na Figura 27.



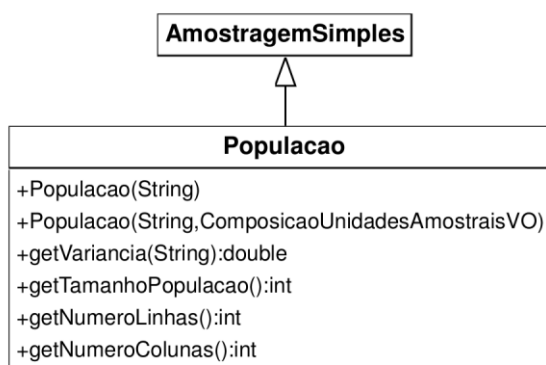


FIGURA 27 – CLASSE QUE REPRESENTA A POPULAÇÃO  
FONTE: O AUTOR (2014)

### Amostragens estratificadas

Para modelar esse tipo de amostragem foi necessário criar uma classe que representasse um estrato dentro de uma amostra estratificada. A classe *EstratoSimples* (Figura 28) é uma especialização da classe *AmostragemSimples*, portanto herda todas as suas características (métodos e atributos).

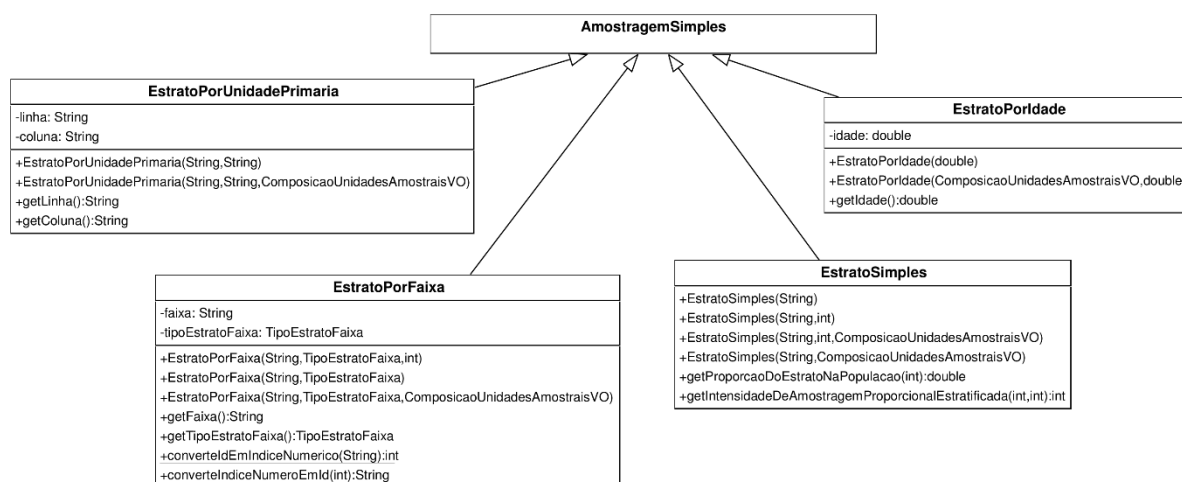


FIGURA 28 – CLASSES QUE REPRESENTAM OS ESTRATOS E SUAS RELAÇÕES  
FONTE: O AUTOR (2014)

Foram representados quatro tipos de estratos de acordo com o critério de estratificação:

- Estrato simples: pode seguir qualquer critério de estratificação que o programador da biblioteca de programação desejar.

- Estrato por idade: O critério de estratificação é a idade, portanto as unidades amostrais que fazem parte deste estrato são da mesma idade.
- Estrato por faixa: Todas as unidades amostrais devem pertencer a mesma faixa (linha ou coluna), dentro de uma localização geográfica.
- Estrato por unidade primária: Todas as unidades amostrais do estrato devem pertencer à mesma unidade primária, ou seja, formarão uma nova unidade amostral denominada unidade primária.

Foi criada uma classe genérica e abstrata (*AmostragemEstratificadaAbstrata*) (Figura 29) para representar o que há em comum entre todos os processos de amostragem que utilizam estratégias de estratificação. Por exemplo, os métodos que calculam a média estratificada, a variância estratificada, a variância entre e dentro dos estratos são comuns a esses processos; por isso, foram colocados nessa classe.

A classe *AmostragemEstratificadaAbstrata* é abstrata, pois os processos de amostragem que utilizam métodos de estratificação são representados por classes específicas e herdam as características dessa classe abstrata. Ela é uma classe genérica, pois pode ser utilizada com qualquer tipo de estrato implementado (Figura 28).

Pode-se observar na Figura 29 que a classe *AmostragemEstratificadaAbstrata* tem uma relação de herança com a classe *ComposicaoUnidadesAmostrais*. Essa relação é importante, pois desse modo a classe *AmostragemEstratificadaAbstrata* se comporta como uma composição de unidades amostrais. Nesse caso as unidades amostrais serão os estratos, que também são composições de unidades amostrais. Portanto cria-se uma relação recursiva, onde a classe *AmostragemEstratificadaAbstrata* é uma composição de estratos de unidades amostrais.

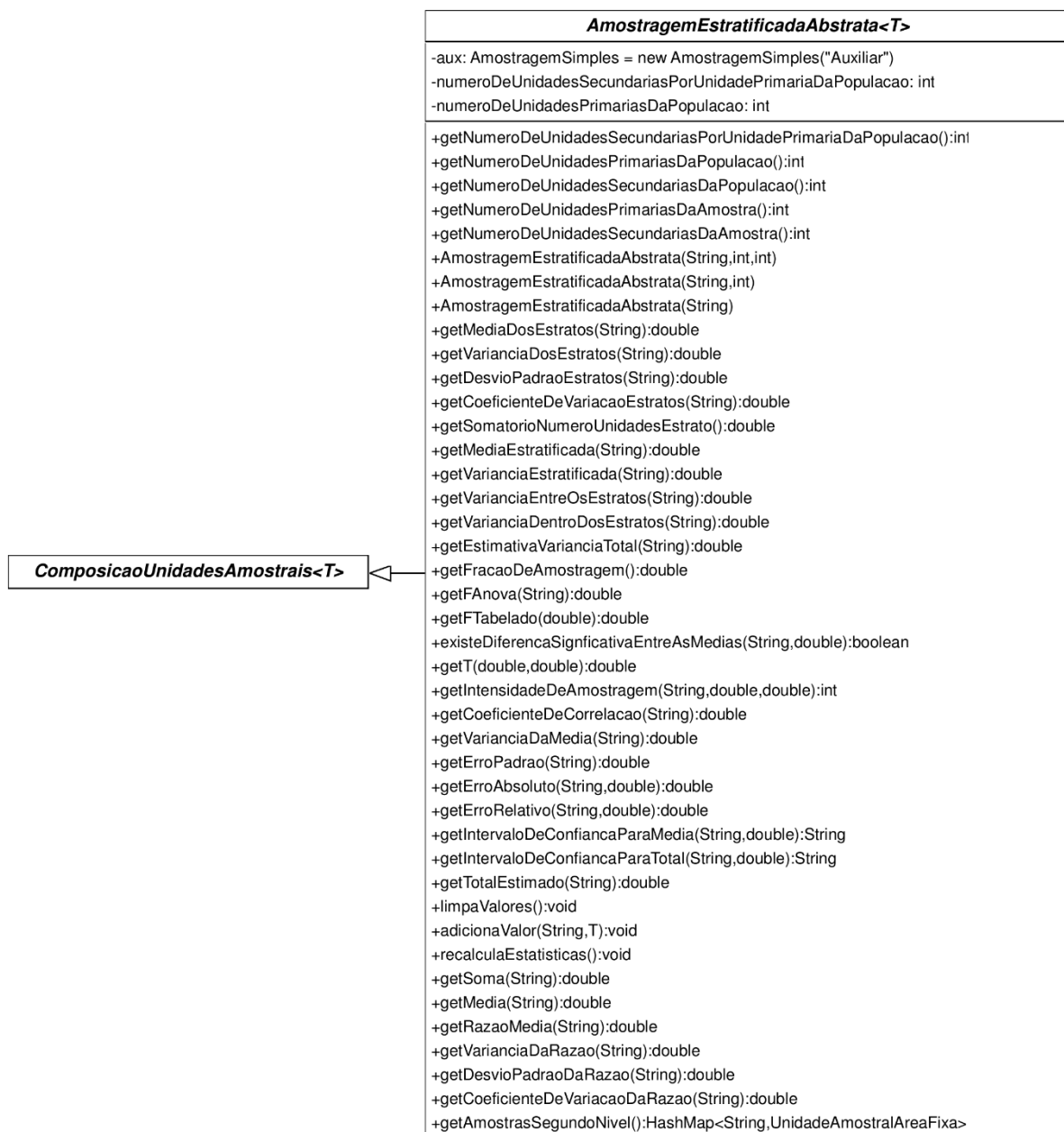


FIGURA 29 – CLASSE QUE REPRESENTA A AMOSTRAGEM ESTRATIFICADA  
 FONTE: O AUTOR (2014)

A biblioteca de programação implementa, além da classe *Populacao* (Figura 27), mais três classes que calculam os parâmetros de uma população: *PopulacaoPorFaixa*, *PopulacaoPorUnidadePrimaria* e *PopulacaoPorIdade* (Figura 30). Enquanto a classe *Populacao* utiliza o censo simples de toda a população para cálculo dos parâmetros, as classes citadas utilizam a estratificação e podem calcular os parâmetros para cada estrato.

A Figura 30 ilustra as classes *PopulacaoPorFaixa*, *PopulacaoPorUnidadePrimaria* e *PopulacaoPorIdade* e suas relações. Pode-se observar que as três classes têm uma relação de herança com a classe *AmostragemEstratificadaAbstrata*. Essa relação determina que as classes tenham os métodos de cálculo da classe *AmostragemEstratificadaAbstrata*, mas com suas respectivas particularidades.

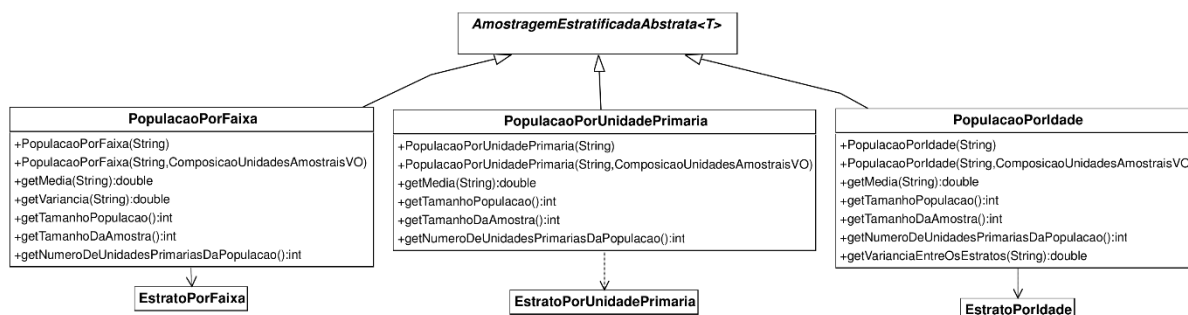


FIGURA 30 – CLASSES QUE REPRESENTAM A POPULAÇÃO POR FAIXA, A POPULAÇÃO POR UNIDADE PRIMÁRIA E A POPULAÇÃO POR IDADE  
 FONTE: O AUTOR (2014)

As classes *EstratoPorFaixa*, *EstratoPorUnidadePrimaria* e *EstratoPorIdade* são utilizadas respectivamente pelas classes *PopulacaoPorFaixa*, *PopulacaoPorUnidadePrimaria* e *PopulacaoPorIdade* para determinar o tipo da estratificação. Esse tipo genérico é determinado quando a respectiva classe herda da classe *AmostragemEstratificadaAbstrata*. Por exemplo, pode-se afirmar que a classe *PopulacaoPorFaixa* é uma *AmostragemEstratificadaAbstrata* do tipo *EstratoPorFaixa*. Com isso a biblioteca de programação ganha flexibilidade e pode ser estendida com outras classes que utilizem métodos estratificados.

A classe *AmostragemEstratificada* (Figura 31) representa uma estratificação simples da população por um critério qualquer. Por esse motivo utiliza a classe *EstratoSimples* para tipificar a estratificação. A classe *AmostragemEstratificada* herda as características da *AmostragemEstratificadaAbstrata*, tipificada por *EstratoSimples* e implementa suas próprias características, tais como: cálculo do erro absoluto considerando o número efetivo de graus de liberdade. Portanto, pode-se afirmar que

a classe *AmostragemEstratificada* é uma *AmostragemEstratificadaAbstrata* do tipo *EstratoSimples*.

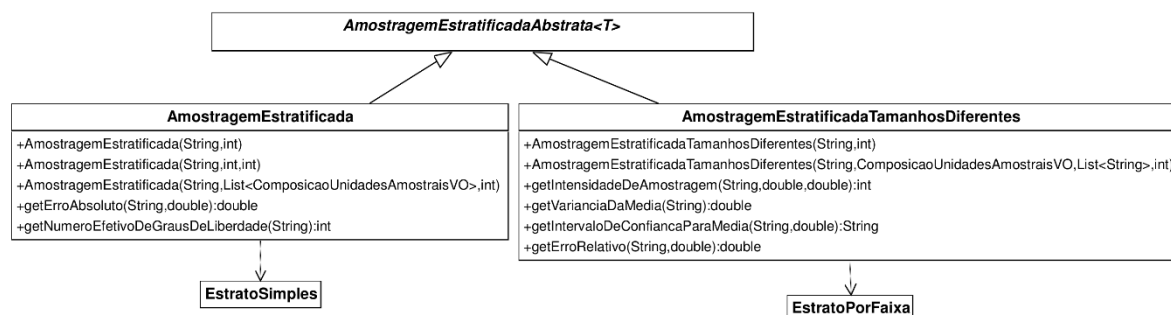


FIGURA 31 – CLASSES QUE REPRESENTAM A AMOSTRAGEM ESTRATIFICADA E A AMOSTRAGEM ESTRATIFICADA DE TAMANHOS DIFERENTES  
FONTE: O AUTOR (2014)

Nessa biblioteca de programação foi implementada a classe *AmostragemEstratificadaTamanhosDiferentes* (Figura 31), que também utiliza a estratégia da estratificação, mas possibilita o uso de estratos com unidades amostrais de tamanhos diferentes. Observa-se que essa classe herda as características da classe *AmostragemEstratificadaAbstrata*, por outro lado implementa as suas particularidades, como o cálculo da intensidade de amostragem, variância da média, intervalo de confiança para a média e erro relativo.

### Amostragem em múltiplos estágios

A amostragem em múltiplos estágios foi representada nessa biblioteca de programação pela classe *AmostragemEmDoisEstágios* (Figura 32), que herda algumas características da classe *AmostragemEstratificadaAbstrata* e implementa os cálculos que são diferentes, tais como: intensidade de amostragem para as unidades primárias, intensidade de amostragem, variância da média, média e média estratificada. Pode-se afirmar, observando a Figura 32, que essa classe utiliza a classe *EstratoPorUnidadePrimaria* para tipificar a amostragem estratificada. Afirma-se, portanto, que a classe *AmostragemEmDoisEstagios* é uma *AmostragemEstratificadaAbstrata* do tipo *EstratoPorUnidadePrimaria*.

A amostragem em conglomerados é definida por Péllico e Brena (1997, p. 201) e Sanquetta *et al.* (2009, p. 134) como uma variação da amostragem em dois estágios, onde o segundo estágio é organizado deterministicamente. Essa definição justifica a criação de uma relação de herança entre as classes *AmostragemEmConglomerados* (Figura 32) e *AmostragemEmDoisEstagios* ilustrada na Figura 32, uma vez que a classe *AmostragemEmConglomerados* é uma especialização da classe *AmostragemEmDoisEstagios*, com a implementação de suas características particulares, tais como: variância da média e intensidade de amostragem.

Queiroz (2012, p. 121) afirma que a amostragem com múltiplos inícios aleatórios é similar à da amostragem em conglomerados. Péllico e Brena (1997, p. 225) complementam que a amostragem com múltiplos inícios aleatórios possui um conglomerado apenas, na qual a unidade conglomerada consiste de um número de subunidades distribuídas uniformemente sobre a população. Por esse motivo utilizou-se uma relação de herança entre as classes *AmostragemEmConglomerados* e *AmostragemSistemáticaComMúltiplosIníciosAleatórios* (Figura 32).

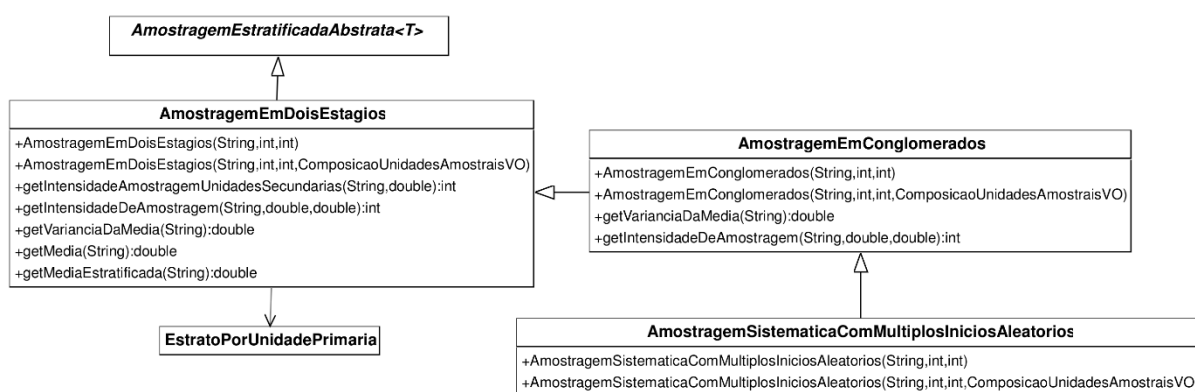


FIGURA 32 – CLASSES QUE REPRESENTAM A AMSTRAGEM EM DOIS ESTÁGIOS, A AMOSTRAGEM EM CONGLOMERDADOS E A AMOSTRAGEM SISTEMÁTICA COM MÚLTIPLOS INÍCIOS ALETÓRIOS  
FONTE: O AUTOR (2014)

### Amostragens em sucessivas ocasiões

Os processos de amostragens em sucessivas ocasiões são utilizados nos inventários florestais contínuos. Esses processos foram revisados na seção 2.1.8.

A classe *AmostragemMultiplasOcasioes* (Figura 33) é abstrata e uma generalização da amostragem em sucessivas ocasiões. Essa classe possui os métodos e cálculos estatísticos que são comuns entre os quatro tipos de amostragem em múltiplas ocasiões implementadas. Os atributos *primeiraOcasiao* e *segundaOcasiao* são do tipo *AmostragemSimples*, e todas as estimativas estatísticas reaproveitam os métodos implementados nessa classe.

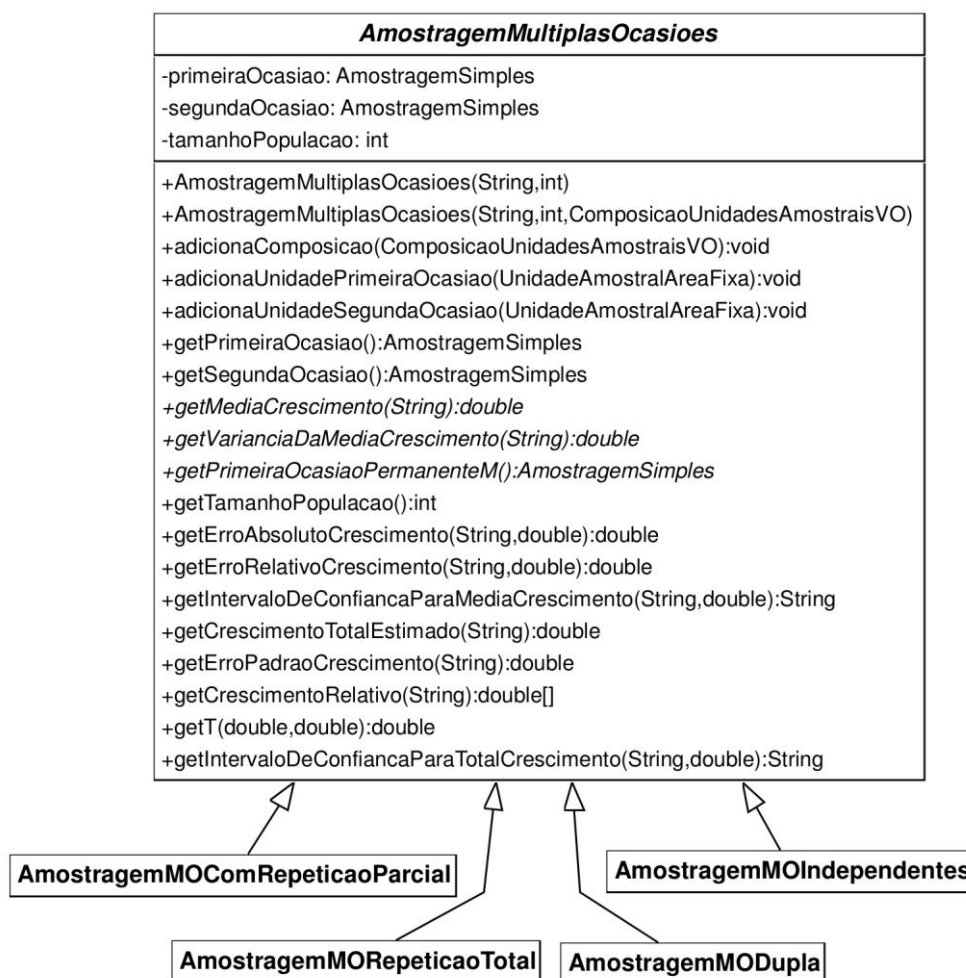


FIGURA 33 – CLASSES QUE REPRESENTAM A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES  
FONTE: O AUTOR (2014)

Os quatro processos de amostragem em sucessivas ocasiões implementados por essa biblioteca de programação são subclasses da classe *AmostragemMultiplasOcasioes* que estão descritas a seguir.

### Sucessivas ocasiões: Amostragens independentes (AI)

A primeira e a segunda ocasiões são tomadas de forma independentes, portanto todas as unidades amostrais são temporárias (PÉLLICO NETTO e BRENA, 1997, p. 236).

Na Figura 34 está apresentada a classe *AmostragemMOIndependentes* que implementa esse tipo de amostragem. Essa classe implementa os métodos *getMediaCrescimento*, *getVariancaDaMediaEstratificada* e *getPrimeiraOcasiao PermanenteM* que são abstratos na classe herdada e possuem implementações específicas para amostragens independentes. Todos os outros métodos que calculam as estimativas para o crescimento, primeira e segunda ocasiões, estão implementados na classe herdada (*AmostragemMultiplasOcasioes*).

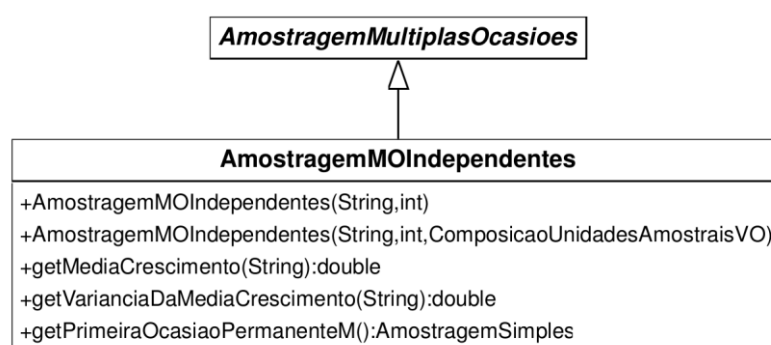


FIGURA 34 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES INDEPENDENTES  
FONTO: O AUTOR (2014)

### Sucessivas ocasiões: Amostragens com repetição total (ART)

As unidades amostrais tomadas na primeira ocasião são remeidas na segunda ocasião, bem como em todas as ocasiões sucessivas (PÉLLICO NETTO e BRENA, 1997, p. 236).

A Figura 35 apresenta a classe *AmostragemMOREpeticaoTotal* que implementa esse tipo de amostragem. Os métodos implementados por essa classe são os métodos abstratos da classe *AmostragemMultiplasOcasioes* e que calculam as estimativas particulares das amostragens com repetição total.



O método *getCrescimentoRelativo* foi reescrito na classe *AmostragemMORepeticaoTotal*, pois considera que o número de unidades amostrais da primeira e da segunda ocasiões devem ser iguais. Caso sejam diferentes a biblioteca de programação não realiza o cálculo e lança uma exceção.

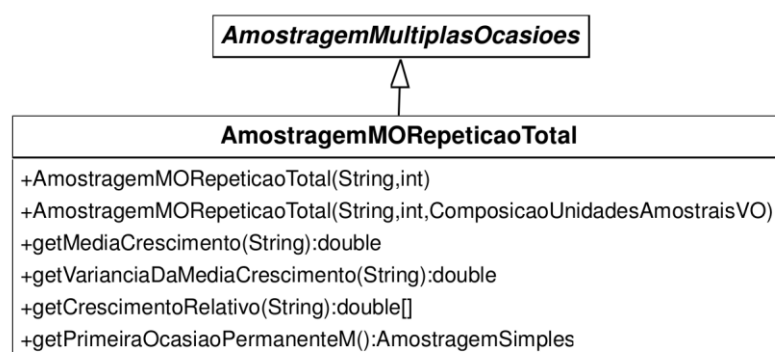


FIGURA 35 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO TOTAL  
 FONTE: O AUTOR (2014)

### Sucessivas ocasiões: Amostragem dupla (AD)

Nesse tipo de amostragem apenas uma parte das unidades amostrais tomadas na primeira ocasião é remeida na segunda ocasião (PÉLLICO NETTO e BRENA, 1997, p. 237).

A classe *AmostragemMODupla* (Figura 36) possui alguns atributos para representar as subamostras:

- *primeiraOcasiaoN1*: conjunto de unidades amostrais que representam todas as amostras da primeira ocasião.
- *primeiraOcasiaoU*: conjunto de unidades amostrais que representam a subamostra temporária da primeira ocasião, ou seja, são as unidades amostrais que foram medidas somente na primeira ocasião, sendo um subconjunto de *n1*.
- *primeiraOcasiaoM*: conjunto de unidades amostrais que representam a subamostra permanente da primeira ocasião, ou seja, são as unidades

amostrais que foram medidas na primeira e segunda ocasião, sendo um subconjunto de n1.

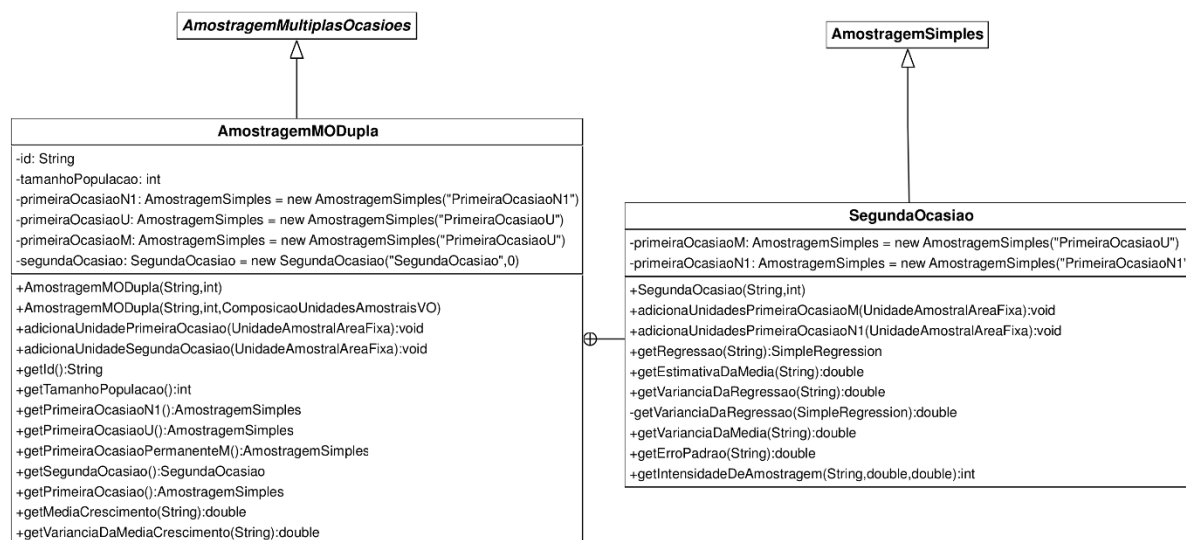


FIGURA 36 – CLASSE QUE REPRESENTA A AMOSTRAGEM DUPLA EM MÚLTIPLAS OCASIÕES  
FONTE: O AUTOR (2014)

A subamostra de unidades amostrais pertencentes à segunda ocasião é representada como uma classe interna da *AmostragemMODupla* chamada *SegundaOcasiao*. Essa classe herda todas as características de *AmostragemSimples*, mas implementa suas particularidades, tais como: o cálculo da intensidade de amostragem para a segunda ocasião e a regressão que estima valores para as unidades amostrais da segunda ocasião que não foram medidas.

### Sucessivas ocasiões: Amostragem com repetição parcial (ARP)

Nesse processo um subconjunto de unidades amostrais da primeira ocasião é remedido na segunda ocasião. Ou seja, na segunda ocasião algumas unidades amostrais da ocasião anterior são remedidas e outras unidades amostrais são acrescentadas (PÉLLICO NETTO e BRENA, 1997, p. 282).

O diagrama da Figura 37 apresenta a classe *AmostragemMOComRepeticaoParcial* e suas classes internas *PrimeiraOcasiaoN1* e *SegundaOcasiao*. Essas classes herdam os métodos e atributos da classe *AmostragemSimples*, mas implementam suas particularidades, tais como: média estimada e variância estimada.

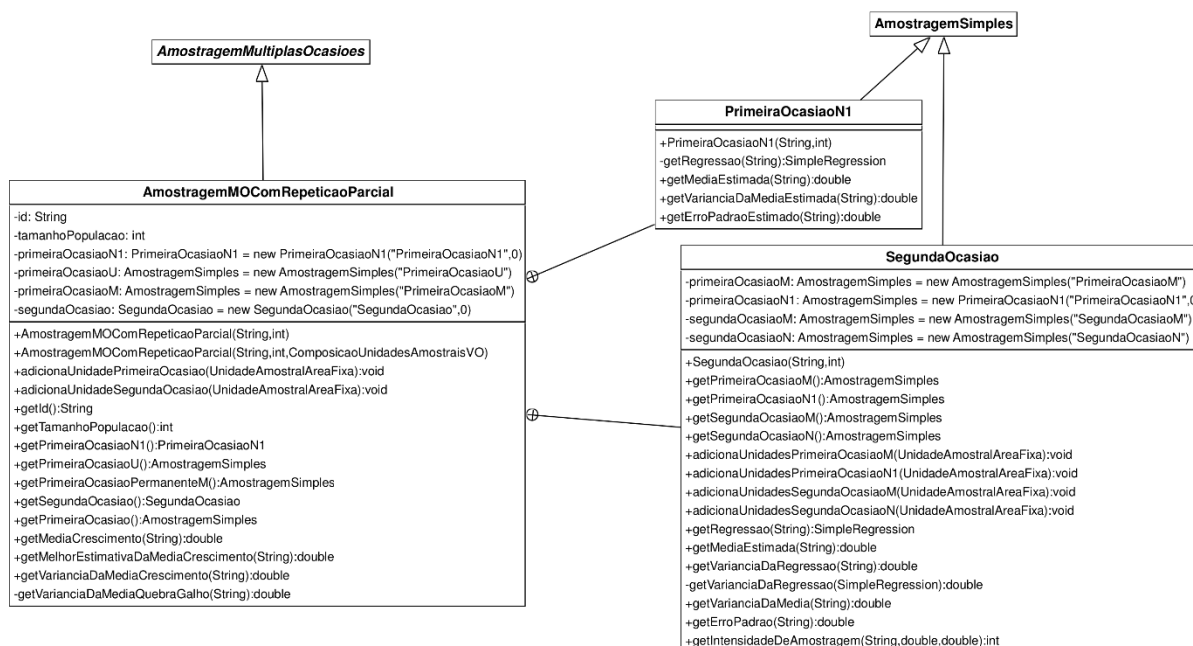


FIGURA 37 – CLASSE QUE REPRESENTA A AMOSTRAGEM EM MÚLTIPLAS OCASIÕES COM REPETIÇÃO PARCIAL  
FONTE: O AUTOR (2014)

Particularmente, a subclasse *SegundaOcasiao* representa o conjunto de unidades amostrais da segunda ocasião, considerando-se permanentes e temporárias. A estrutura dessa subclasse armazena em seus atributos, além da *primeiraOcasiaoN1* e *primeiraOcasiaoM*, a representação das seguintes subamostras:

- *segundaOcasiaoM*: conjunto de valores medidos na segunda ocasião de unidades amostrais permanentes.
- *segundaOcasiaoN*: conjunto de valores medidos na segunda ocasião de unidades amostrais novas, ou seja, que não existiam na primeira ocasião.

A subclasse *PrimeiraOcasiaoN1* representa o conjunto de unidades amostrais da primeira ocasião, considerando-se permanentes e temporárias. Essa subclasse foi criada para calcular as estimativas da média, variância e erro padrão, considerando-se a regressão linear que estima valores na primeira e segunda ocasiões, no caso de unidades amostrais que não são permanentes. O método que calcula a média estimada (*getMediaEstimada*) é utilizado no cálculo da melhor estimativa da média para o crescimento (*getMelhorEstimativaDaMediaCrescimento*).

#### 4.1.5 Resultados dos testes da biblioteca de programação

Os testes da biblioteca de programação, foram realizados de duas maneiras:

- Foram resolvidos, com a biblioteca de programação, todos os exercícios propostos no livro de Péllico Netto e Brena (1997, p. 97-300) e apresentados os resultados em uma interface de texto, conforme apresentado no Apêndice 5.
- Foram realizados testes automatizados ao longo do projeto, por meio de testes unitários programados no ambiente de desenvolvimento “Eclipse”. Dois exemplos de evidências destes testes estão apresentados no Apêndice 5.

#### 4.2 APRESENTAÇÃO DO GERADOR DE DADOS PSEUDOALEATÓRIO

Com o objetivo de exemplificar o resultado do gerador de dados pseudoaleatórios, serão apresentados, nesta seção, cada um dos 4 passos (ver seção 3.3) para a geração dos dados do “Aluno 1” referente ao primeiro exercício (desafio “Parâmetros da População”) implementado. Os dados gerados para o primeiro exercício são utilizados também nos exercícios 2, 3, 4, 5, 6, 7, 8 e 9 (ver Quadro 26). Conforme discutido na seção 3.3, para os demais exercícios (de 10 a 16) utilizam-se os mesmos passos para geração dos dados, modificando-se os dados de origem e algumas propriedades da estrutura da população.

##### **Passo 1 – Carga dos dados**

Os dados apresentados no Anexo 1 foram cadastrados no banco de dados antes do início da geração pseudoaleatória para qualquer usuário do sistema.

### **Passo 2 – Criação da semente**

O código de matrícula do “Aluno 1” era PRR20144001<sup>2</sup>. Para se obter a semente (ver definição de semente na seção 2.3), retirou-se as três letras do código, ficando apenas com o valor numérico: “20144001”.

### **Passo 3 – Sorteio da estrutura da população**

No quadro abaixo estão apresentados os valores sorteados para as propriedades da estrutura da população para o “Aluno 1”.

<b>Propriedade</b>	<b>Valor</b>
Quantidade de linhas secundárias	35
Quantidade de colunas secundárias	18
Comprimento da unidade amostral	25 metros
Largura da unidade amostral	30 metros
Quantidade de linhas primárias	7
Quantidade de colunas primárias	3

QUADRO 25 – PRPOPIEDADES DA ESTRUTURA DA POPULAÇÃO PARA O “ALUNO 1  
FONTE: O AUTOR (2014)”

### **Passo 4 – Geração pseudoaleatória dos dados estratificados por idade**

A quantidade de dados gerada para o “Aluno 1”, estratificada por idade foi: 222 para 6 anos; 216 para 9 anos e 192 para 12 anos. Totalizando 630 unidades amostrais de área fixa de 25 metros de comprimento e 30 metros de largura, estruturada em 35 linhas e 18 colunas. As 630 unidades de área fixa foram ainda subdividas em 21 unidades primárias. Os dados gerados neste exemplo estão apresentados no Apêndice 6.

### **Testes dos dados gerados com Kolmogorov-Smirnov**

Esta subseção apresenta um exemplo de como foi realizado o teste de Kolmogorov-Smirnov com o objetivo de comprovar que os dados gerados pelo método apresentado na seção 3.3 podem ser considerados da mesma distribuição de probabilidade do que os dados de origem.

---

<sup>2</sup> O código da matrícula apresentado é fictício, apesar de seguir o formato de código de matrícula utilizado pela UFPR.

Conforme informado no passo 4 da seção 3.3, a geração dos dados foi realizada para cada conjunto de dados estratificados por idade, portanto, o teste de Kolmogorov-Smirnov também foi realizado para cada estrato gerado.

O primeiro passo foi obter o histograma separado por idade, e como consequência o tamanho das classes para os dados originais (Anexo 1). A Figura 1 ilustra os três histogramas obtidos, bem como o tamanho das classes para cada estrato por idade.

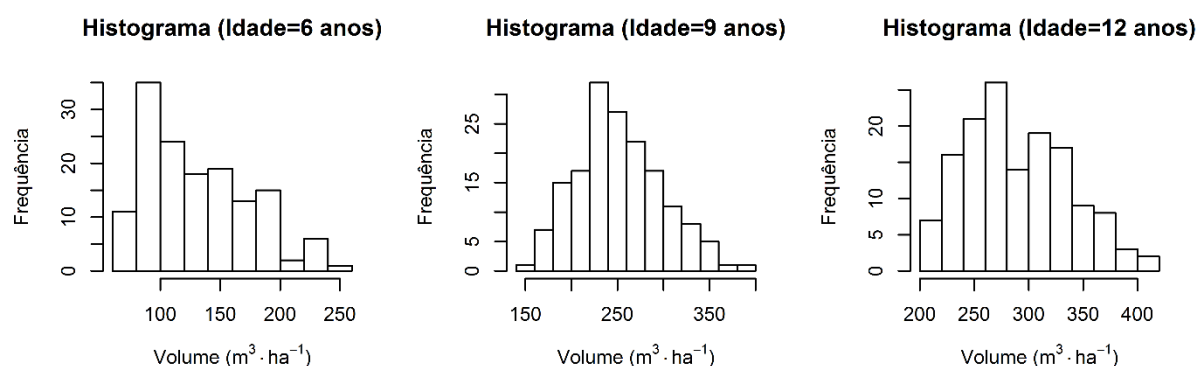


FIGURA 38 – HISTOGRAMA DOS DADOS ORIGINAIS (ANEXO 1) ESTRATIFICADOS POR IDADE  
 FONTE: O AUTOR(2014)

Pode-se observar pelos histogramas da Figura 38 que o tamanho da classe é de  $20\text{m}^3/\text{ha}$  para todas as idades, variando o número classes: 10 para 6 anos, 13 para 9 anos e 11 para 12 anos.

Para o cálculo do valor  $d_{cal}$  do teste de Kolmorov-Smirnov foi necessário, além das frequências absoltas (presente no histograma) obter as frequências relativas e acumuladas para cada idade. Para exemplificar, a Tabela 4 demonstra esse cálculo para os dados com idade de 6 anos do Anexo 1 (dados originais) e do Apêndice 6 (dados gerados).

Por motivo de apresentação, os títulos das colunas da Tabela 4 foram abreviadas, sendo que seus significados são: F – Frequência, Fre – Frequência relativa, Fac – Frequência relativa acumulada, Ori – Dados Originais (Anexo 1), Ger – Dados gerados (Apêndice 6).

TABELA 1 – EXEMPLO DO TESTE DE KOLMOGOROV-SMIRNOV PARA OS DADOS GERADOS  
 FONTE: O AUTOR (2014)

Classes	F (Ori)	F (Ger)	Fre (Ori)	Fre (Ger)	Fac (Ori)	Fac (Ger)	Diferença
60-80	11	21	0,076	0,095	0,076	0,095	0,018
80-100	35	53	0,243	0,239	0,319	0,333	0,014
100-120	24	34	0,167	0,153	0,486	0,486	0,000
120-140	18	34	0,125	0,153	0,611	0,640	0,029
140-160	19	29	0,132	0,131	0,743	0,770	0,027
160-180	13	15	0,090	0,068	0,833	0,838	0,005
180-200	15	21	0,104	0,095	0,938	0,932	0,005
200-220	2	5	0,014	0,023	0,951	0,955	0,004
220-240	6	10	0,042	0,045	0,993	1,000	0,007
240	1	0	0,007	0,000	1,000	1,000	0,000
Soma	144	222				$d_{cal}$	0,0285

O cálculo do valor crítico, ou tabelado ( $d_{tab}$ ), foi realizado de acordo com a equação 17 apresentada na seção 3.3, com os valores  $m=144$  (número de unidades amostrais dos dados originais para idade = 6 anos) e  $n=222$  (número de unidades amostrais geradas para idade = 6 anos). A equação 18 demonstra o cálculo do valor crítico para este caso.

$$d_{tab} = 1,36 \cdot \sqrt{\frac{144 + 222}{144 \cdot 222}} = 0,1455 \quad (18)$$

Nesse caso, o valor  $d_{cal}=0,0285$  é menor do que o valor  $d_{tab}=0,1455$ , portanto não se tem evidências para rejeitar a hipótese da nulidade de que os dois conjuntos de dados, original (Anexo 1) e gerado (Apêndice 6), tenham em comum a mesma distribuição de dados.

Para comparar os dois conjuntos de dados testados nesse exemplo, a Figura 39 ilustra o gráfico da função acumulada.

Para se testar os outros conjuntos de dados (idade de 9 e 12 anos) para esse caso, procede-se os mesmos passos demonstrados nesta seção. O Apêndice 7 demonstra os testes realizados para 102 conjuntos de dados, totalizando os dados de 34 alunos do sistema. Pode-se observar no Apêndice 7 que em todos os 102 conjuntos de dados gerados, o valor calculado ( $d_{cal}$ ) é menor do que o valor crítico ( $d_{tab}$ ), portanto não se tem evidências para rejeitar a hipótese da nulidade de que os conjuntos de dados gerados e originais, tenham em comum a mesma distribuição de dados.

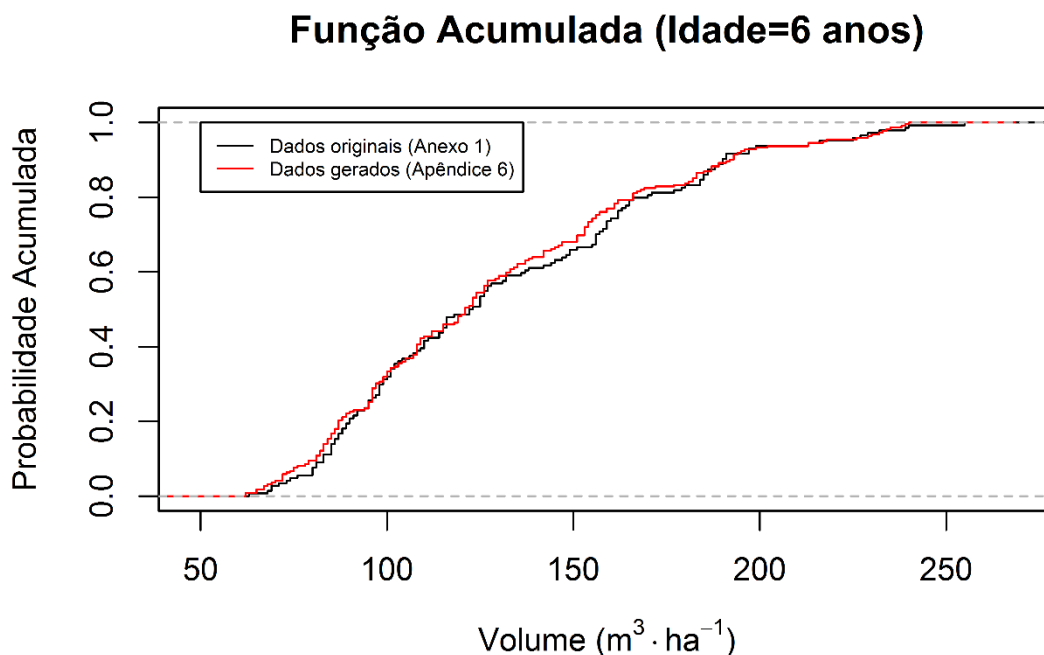


FIGURA 39 – COMPARATIVO DOS DADOS ORIGINAIS (ANEXO 1) E DADOS GERADOS (APÊNDICE 6) POR MEIO DA FUNÇÃO ACUMULADA DE PROBABILIDADE  
 FONTE: O AUTOR (2014)

#### 4.3 APRESENTAÇÃO DO SISTEMA

O sistema de aprendizagem desenvolvido possui um esquema básico de funcionamento, ilustrado na Figura 40.

O professor deve fornecer os conceitos teóricos sobre os processos de amostragem, pois o *software* não os contém disponibilizados para utilização em inventários florestais. O aluno se cadastra no sistema e tem acesso ao conjunto de exercícios propostos. Os dados e as estruturas das unidades amostrais são criados pelo sistema de maneira aleatória por meio de um gerador e, como consequência, os alunos terão exercícios personalizados. O aluno pode exportar os dados das unidades amostrais para uma planilha eletrônica, efetuar todos os cálculos necessários, retornar ao sistema e conferir os valores calculados por ele. O sistema realiza os cálculos dos exercícios propostos para os processos de amostragem, compara o resultado informado pelo aluno e relata se houve acerto ou erro.



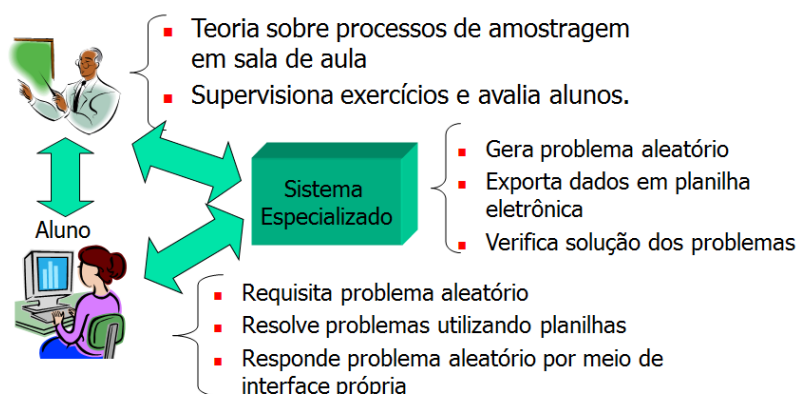


FIGURA 40 – ESQUEMA BÁSICO DE FUNCIONAMENTO  
FONTE: O AUTOR (2014)

As interações entre o sistema desenvolvido e o aluno podem ser ilustradas por meio de um fluxograma básico (Figura 41), no qual o primeiro passo é a criação do *login* do aluno. O segundo passo (executado pelo sistema) é a criação dos 16 desafios (ou exercícios) sobre os processos de amostragem em um inventário florestal. Nesse passo os exercícios são gerados pelo sistema com dados aleatórios e individualizados por aluno. Esse passo foi apresentado em detalhes na seção 4.2. No terceiro passo o aluno realiza o *login* no sistema e tem acesso aos exercícios propostos. O quarto passo é a resolução dos exercícios pelo aluno, que é corrigido pelo sistema no quinto passo. O quarto e o quinto passos são repetidos várias vezes até que o aluno tenha completado um desafio e possa iniciar a resolução do próximo.

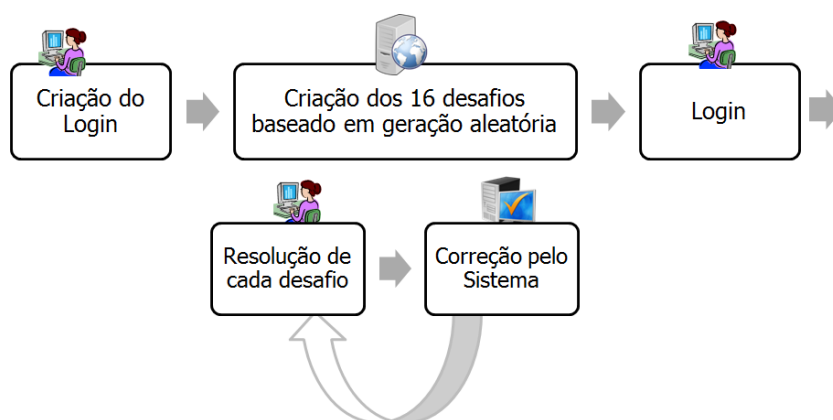


FIGURA 41 – FLUXOGRAMA DE FUNCIONAMENTO  
FONTE: O AUTOR (2014)

Cada desafio implementa um processo de amostragem e é composto por um conjunto de questões que devem ser resolvidas pelo aluno. O Quadro 26 apresenta a quantidade de questões para cada desafio e o total de questões existentes no sistema.

Número	Nome do Desafio (Processos de Amostragem)	Questões
1	Parâmetros da População	5
2	Parâmetros da População Estratificada	5
3	Parâmetros da População por Faixa de Tamanho Igual	5
4	Parâmetros da População por Faixa de Tamanho Diferente	5
5	Parâmetros da População por Unidade Primária	3
6	Amostragem Aleatória Simples	18
7	Amostragem Aleatória Simples - Unidades Amostrais de Tamanhos Diferentes	16
8	Amostragem Estratificada	19
9	Amostragem Sistemática em Dois Estágios	14
10	Amostragem em Dois Estágios	20
11	Amostragem em Conglomerados	20
12	Amostragem Sistemática com Múltiplos Inícios Aleatórios	20
13	Amostragem em Múltiplas Ocasões - Amostras Independentes	28
14	Amostragem em Múltiplas Ocasões - Amostras Com Repetição Total	28
15	Amostragem em Múltiplas Ocasões - Amostragem Dupla	36
16	Amostragem em Múltiplas Ocasões - Repetição Parcial	36
	Total	278

QUADRO 26 – DESAFIOS E QUESTÕES  
FONTE: O AUTOR (2014)

Como o *software* está disponível na internet, basta que o aluno digite a URL (*Uniform Resource Locator*) do sistema em um navegador para que tenha acesso à tela de entrada (ilustrada na Figura 42a).

Nessa tela (Figura 42a) o aluno pode entrar no sistema, preenchendo os dados de matrícula, senha, e clicando no botão “Entrar”. Também é possível criar um novo usuário por meio do botão no canto inferior esquerdo. Se optar por criar um novo usuário, então a tela ilustrada na Figura 42b será mostrada. Quando um novo usuário é criado, o sistema inclui os 16 desafios para ele. No contexto desse *software*, desafios se referem aos exercícios que o aluno deverá resolver. No momento da criação de um novo usuário, o sistema sorteia valores para as unidades amostrais de cada desafio; com isso, cada aluno possuirá o seu conjunto de dados para os exercícios. O método utilizado para a criação dos dados foi o núcleo estimador variável (ver seção 3.3).

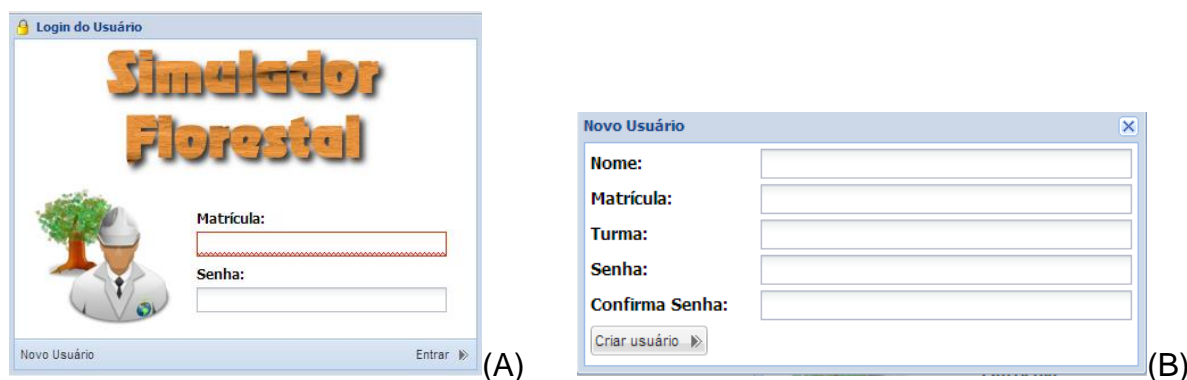


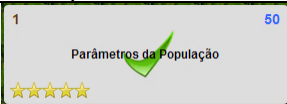



FIGURA 42 – (A) TELA DE LOGIN (B) TELA DE CRIAÇÃO DE NOVO USUÁRIO  
FONTE: O AUTOR (2014)

Os desafios criados pelo *software* são exercícios que cobrem 16 processos de amostragem (listados no Quadro 26). Após o *login* (por meio da tela ilustrada na Figura 42a), o aluno é direcionado para uma tela que apresenta o estado dos desafios (Figura 43).



FIGURA 43 – TELA INICIAL  
FONTE: O AUTOR (2014)

Os exercícios podem assumir quatro estados (Finalizado, Disponível, Iniciado e Indisponível), que são indicados pelos ícones e pela transparência do quadro do desafio. O Quadro 27 apresenta os quatro estados, sendo que na primeira coluna está apresentado o nome do estado, na segunda coluna a descrição do que o estado representa e na terceira coluna um exemplo de desafio no estado em questão.

Estado	Descrição	Exemplo
Finalizado	Representa um desafio que o aluno terminou, ou seja, todas as questões foram respondidas corretamente.	
Disponível	Representa um desafio que o aluno pode iniciar. O aluno ainda não respondeu nenhuma questão nesse desafio.	
Iniciado	Representa um desafio que está disponível, mas o aluno já iniciou a sua resolução, com respostas corretas ou incorretas.	
Indisponível	Representa um desafio que ainda não pode ser iniciado, pois é necessário que o aluno finalize o desafio anterior.	

QUADRO 27 – ESTADOS DOS DESAFIOS  
FONTE: O AUTOR (2014)

No exemplo da Figura 43, o desafio 1 (Parâmetros da População) foi finalizado, o desafio 2 (Parâmetros da População Estratificada) está disponível para o aluno, mas ainda não foi iniciado e os desafios 3 a 16 estão bloqueados e ainda não podem ser respondidos. Neste exemplo, o desafio 3 se tornará disponível quando o aluno acertar todas as questões do desafio 2.

Cada um dos 16 desafios possui telas específicas. Para se ilustrar o funcionamento do *software* podem-se separar os desafios em grupos que possuem telas similares. A classificação demonstrada no Quadro 28 não tem a intenção de propor uma taxonomia para os processos de amostragem, mas sim de organizar a apresentação do *software* neste artigo.

Grupo	Processos de Amostragem
Grupo 1	Parâmetros da população, parâmetros da população estratificada, parâmetros da população por faixa de tamanho igual, parâmetros da população por faixa de tamanho diferente e parâmetros da população por unidade primária.
Grupo 2	Amostragem aleatória simples, amostragem aleatória simples com unidades amostrais de tamanhos diferentes, amostragem estratificada e amostragem sistemática em dois estágios.
Grupo 3	Amostragem em dois estágios, amostragem em conglomerados e amostragem sistemática com múltiplos inícios aleatórios.
Grupo 4	Amostragem em múltiplas ocasiões com amostras independentes, amostragem em múltiplas ocasiões com repetição total, amostragem em múltiplas ocasiões por amostragem dupla e amostragem em múltiplas ocasiões com repetição parcial.

QUADRO 28 – ORGANIZAÇÃO DOS PROCESSOS DE AMOSTRAGEM  
FONTE: O AUTOR (2014)

Os passos apresentados pelo sistema em cada um dos grupos de desafios estão apresentados a seguir.

### 4.3.1 Desafios do grupo 1

A seguir, estão descritos os passos que o sistema apresenta ao aluno no grupo de desafios 1 para que possa resolver as questões colocadas.

#### Passo 1

A primeira tela desses desafios apresenta as unidades amostrais de uma população de *Pinus sp* organizada em uma tabela com unidades primárias e secundárias, com valores em volume de madeira por hectare, conforme ilustrado na Figura 44.

# Simulador Florestal: Processos de Amostragem

2º - Parâmetros da População Estratificada

Ao lado está representada a área de uma população de *Pinus sp* com **76,8** ha, enumerada totalmente, por meio da divisão em **512** unidades amostrais de forma retangular, com **30m** de largura por **50m** de comprimento, ou seja **1.500 m² (0,15 ha)** de superfície. Em cada uma das unidades amostrais, são representados os volumes com casca, por hectare. A população foi estruturada em linhas, numeradas de (1) a (32), e colunas, identificadas pelas letras (a) até (p), para facilitar a identificação das unidades.

A tabela ao lado representa um povoamento que foi plantado em três épocas distintas. As idades podem ser identificadas pela cor verde de cada unidade amostral. Quanto mais escuro o verde da unidade amostral, mais idade possui o povoamento. Pode-se perceber **3** intensidades de verde. A intensidade mais clara representa um plantio com **6 anos** de idade, a intensidade intermediária representa um plantio com **9 anos** de idade e a intensidade mais escura representa um plantio com **12 anos** de idade. Com base nestas idades do povoamento, a população

Desafios

Exportar Amostras

Amostras

Responder

		I				II				III			
		a	b	c	d	e	f	g	h	i	j	k	
A	1	123	158	141	123	219	109	181	115	93	133	193	
	2	87	164	100	164	83	135	190	199	184	121	197	
	3	86	81	166	124	121	100	125	145	97	140	135	
	4	182	184	87	163	92	112	158	90	98	159	153	
B	5	83	213	96	95	124	107	134	120	157	85	82	
	6	152	185	116	114	167	83	114	182	155	145	100	
	7	114	87	109	139	148	84	82	104	101	189	110	
	8	128	104	157	91	150	138	92	97	86	131	240	
C	9	95	128	112	128	95	183	111	160	183	185	93	
	10	147	118	63	68	137	97	137	142	90	108	98	
	11	101	97	86	230	126	134	124	99	151	177	247	
	12	305	239	217	259	323	206	386	194	275	289	197	
D	13	320	313	308	244	327	301	160	248	258	235	195	
	14	250	270	233	265	260	257	193	230	180	307	306	
	15	268	215	223	287	277	282	284	386	279	188	192	
	16	258	221	281	280	201	289	234	245	348	277	283	
E	17	238	331	271	227	222	277	263	290	200	216	250	
	18	190	182	223	206	234	236	231	205	285	245	245	

Pontuação

ALUNO 40

Sair

FIGURA 44 – PASSO 1 DO DESAFIO 2 (PARÂMETROS DA POPULAÇÃO ESTRATIFICADA)  
FONTE: O AUTOR (2014)

Nesta tela, o aluno pode obter uma planilha em formato xls (Excel) para realizar os seus cálculos referentes a cada processo. Clicando em “Responder”, o



aluno irá para o segundo passo do desafio. Pode-se observar, na Figura 44, que a área da esquerda da tela é reservada para instruções; por elas, o aluno obtém informações que podem ser utilizadas para responder às questões e também orientações para uso do *software*. Todas as telas dos desafios possuem essa área de instruções que é específica para cada uma. Nas próximas figuras essa área não será mostrada.

## Passo 2

No passo 2 o aluno deve responder às questões do desafio referentes aos parâmetros da população representada pelas unidades amostrais da tabela obtida no passo anterior. Na Figura 45 pode-se observar que o aluno respondeu aos parâmetros dos estratos e à média estratificada corretamente, por esse motivo esses campos estão em verde. O parâmetro variância estratificada foi respondido com erro, portanto o sistema apresenta na cor vermelha. Para responder cada um dos parâmetros o aluno deve clicar no ícone ao lado da caixa de texto, ou então no botão “Responder” correspondente, ou ainda no botão “Responder Todas”. Ao fazer isso o sistema avalia se a resposta inserida pelo aluno está correta ou não, informando-o pelas cores verde (caso correto) ou vermelha (caso errado).

A interface do software apresenta uma barra superior com 'Desafios' e 'Exportar Amostras'. Abaixo, uma tabela com 7 colunas: Estrato ( $h$ ), Número Unidades ( $N_h$ ), Média ( $X_h$ ), Variância ( $S^2_h$ ), Desvio Padrão ( $S_h$ ), Coef. de Variação ( $CV_h$ ) e Volume ( $m^2$ ). Os dados são os seguintes:

Estrato $h$	Número Unidades $N_h$	Média $X_h$	Variância $S^2_h$	Desvio Padrão $S_h$	Coef. de Variação $CV_h$	Volume $m^2$
E1	170	19,738	38,33	6,191	31,366	3.355,5
E2	179	38,236	43,189	6,572	17,188	6.844,2
E3	163	43,664	49,676	7,048	16,142	7.117,2

Abaixo da tabela, há um botão 'Responder'. A seção 'Outros Parâmetros' contém:

- Média estratificada: 33,822  $m^2 / 0,15 ha$  (verde)
- Variância estratificada: 43,555  $(m^2 / 0,15 ha)^2$  (vermelha)
- Variância entre os estratos:   $(m^2 / 0,15 ha)^2$
- Variância dentro dos estratos:   $(m^2 / 0,15 ha)^2$

Na base da seção, há botões 'Responder' e 'Responder todas'. A barra inferior indica 'ALUNO 40' e um botão 'Sair'.

FIGURA 45 – PASSO 2 DO DESAFIO 2 (PARÂMETROS DA POPULAÇÃO ESTRATIFICADA)  
FONTE: O AUTOR (2014)

A precisão avaliada pelo software é de três casas decimais. Nos casos em que o dígito à direita (após a vírgula) for zero é possível incluir a resposta com menos do que três casas decimais. Por exemplo, se a resposta correta for 7,820 (considerando arredondamento), então o sistema aceitará como resposta certa o número com 2 casas decimais. Esse comportamento é válido para todos os outros desafios.

#### 4.3.2 Desafios do grupo 2

Os desafios do grupo 2 são compostos pelos passos apresentados nas seções a seguir.

##### **Passo 1**

No primeiro passo desses desafios, o sistema apresenta uma tela com uma tabela sem valores, que representa uma área de uma população de *Pinus sp*, organizada em unidades primárias e secundárias, conforme ilustrada na Figura 46. Nessa tela, o aluno deve selecionar as unidades amostrais do inventário piloto clicando sobre as células, ou sobre as colunas da tabela (dependendo do desafio). No exemplo ilustrado na Figura 46 o aluno escolheu como amostra as seguintes unidades amostrais: 4-c, 3-e, 9-b, 7-e e 7-h. Os valores de volume de madeira por hectare são revelados quando se interage com a tabela. Depois que foram selecionadas as unidades amostrais, o aluno deve clicar em “Próximo”. Péllico Netto e Brena (1997, p. 119) afirmam que o número de unidades do inventário piloto é arbitrado, portanto, no software foi arbitrado que o aluno deve selecionar no mínimo 5 unidades amostrais nessa etapa. O cálculo do número de unidades amostrais para o inventário definitivo é realizado no passo 3 desse desafio.

		I				II					
		a	b	c	d	e	f	g	h	i	j
A	1										
	2										
	3					121					
	4			87							
B	5										
	6										
	7					148			104		
	8										
C	9		128								
	10										
	11										
	12										
	13										

FIGURA 46 – PASSO 1 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES)  
 FONTE: O AUTOR (2014)

## Passo 2

Nesta etapa o aluno visualiza as unidades amostrais escolhidas para o inventário piloto, conforme ilustradas na Figura 47. Ele deve copiar e colar a tabela apresentada para uma planilha com o objetivo de realizar os cálculos do inventário piloto. Clicando em “Próximo” será direcionado para o próximo passo.

Unidade	Coord	Volume
1	3-e	121
2	4-c	87
3	7-e	148
4	7-h	104
5	9-b	128

FIGURA 47 – PASSO 2 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES)  
 FONTE: O AUTOR (2014)

## Passo 3

Neste passo, o aluno deve calcular as estimativas para o inventário piloto com base nas unidades amostrais escolhidas. Após o aluno calcular corretamente essas estimativas, o sistema o direciona para acrescentar as unidades amostrais de acordo com o cálculo da intensidade amostral. Isso é feito na mesma tela demonstrada no



passo 1 (Figura 46). No exemplo ilustrado, na Figura 48, foram selecionadas 5 unidades amostrais, cuja intensidade de amostragem calculada foi de 28 unidades. Por esse motivo, ao clicar em “Próximo” o sistema apresenta a mensagem ilustrada na Figura 48b e direciona o aluno para a tela da Figura 46 com o objetivo de acrescentar as unidades amostrais faltantes para o inventário definitivo.

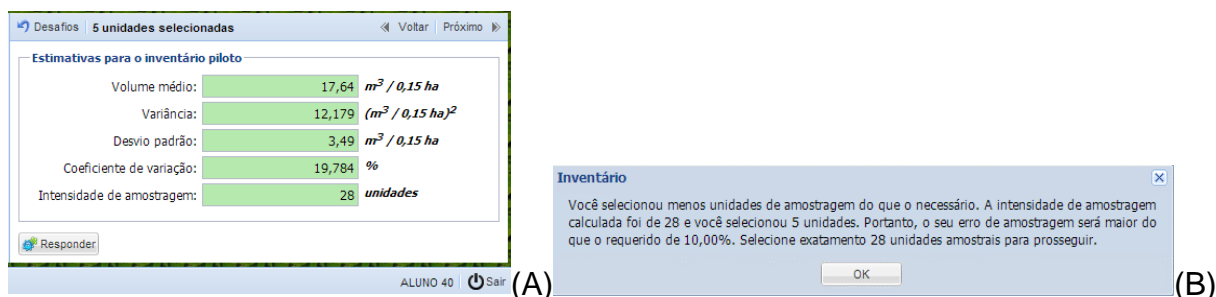


FIGURA 48 – (A) PASSO 3 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES) (B) MENSAGEM DO SISTEMA  
FONTE: O AUTOR (2014)

## Passo 4

No último passo deste grupo de desafios o aluno deve inserir as estimativas para o inventário definitivo. A tabela com as unidades amostrais do inventário definitivo é obtida do mesmo modo que foi demonstrado no passo 2 (Figura 47). A tela ilustrada pela Figura 49 apresenta um exemplo das estimativas para o inventário definitivo que foram calculadas pelo aluno e corrigidas pelo sistema, finalizando assim o desafio.

(A) Screenshot of the 'Desafios' window showing '28 unidades selecionadas'. The table displays the following data:

Estimativas para o inventário definitivo	Valor	Unidade
Média:	38,202	$m^3 / 0,15 \text{ ha}$
Desvio padrão:	11,843	$m^3 / 0,15 \text{ ha}$
Variância:		$(m^3 / 0,15 \text{ ha})^2$
Variância da média:		$(m^3 / 0,15 \text{ ha})^2$
Erro Padrão:		$m^3 / 0,15 \text{ ha}$
Coefficiente de variação:		%
Variância da média relativa:		%
Erro de amostragem absoluto:		$m^3 / 0,15 \text{ ha}$
Erro de amostragem relativo:		%
Intervalo de confiança para a média:	$m^3 / 0,15 \text{ ha} \leq 38,202 \leq$	$m^3 / 0,15 \text{ ha} = 99,00 \%$
Total da população:		$m^3$
Intervalo de confiança para o total:	$m^3 \leq X \leq$	$m^3 = 99,00 \%$
Estimativa mínima de confiança:		$m^3 / 0,15 \text{ ha}$

FIGURA 49 – PASSO 4 DO DESAFIO 6 (AMOSTRAGEM ALEATÓRIA SIMPLES)  
FONTE: O AUTOR (2014)

### 4.3.3 Desafios do grupo 3

O grupo de desafios 3 possui quatro passos que estão apresentados nas seções seguintes.

#### Passo 1

As unidades amostrais são pré-selecionadas pelo sistema. Isso significa que a primeira tela apresenta um conjunto delas, estruturadas de acordo com o tipo de amostragem, conforme ilustrado na Figura 50. Por outro lado, a estrutura tem como padrão a separação em unidades primárias e unidades secundárias. As unidades secundárias sempre pertencem a uma unidade primária e todas as unidades primárias possuem o mesmo número de unidades secundárias.

Desafios							
UNIDADE PRIMÁRIA	UNIDADE SECUNDÁRIA						
	(m <sup>2</sup> /0,25 ha)						
Bloco	I	II	III	IV	V	VI	VII
1	216,337	257,272	228,679	215,612	473,421	185,800	331,814
2	173,595	154,755	212,869	431,059	106,405	686,680	106,372
3	246,081	277,144	113,301	283,464	141,373	283,804	100,802
4	215,628	427,856	378,318	364,211	210,662	434,405	304,359
5	75,878	218,276	210,118	165,610	109,846	96,870	287,672
6	190,397	182,311	170,834	193,621	234,515	162,724	225,613
7	215,427	209,056	104,674	155,744	81,251	397,264	334,515
8	339,653	111,885	281,122	118,623	192,436	257,988	120,845
9	136,916	273,117	249,060	257,287	294,120	686,680	286,999
10	94,461	102,566	452,710	163,392	259,361	184,839	168,057
11	240,982	163,475	255,665	254,355	97,578	296,202	164,007
12	186,442	504,524	249,166	259,141	100,365	151,483	259,763
13	290,299	259,412	197,750	466,774	237,444	150,468	134,216
14	283,778	220,332	281,326	94,120	123,003	194,177	226,778
15	160,374	162,395	288,956	227,573	99,167	84,615	138,001
16	105,260	163,673	167,627	158,003	104,166	296,317	213,799
17	213,753	159,903	246,311	114,442	245,851	132,290	158,671
18	434,989	97,446	128,840	506,591	308,625	270,391	297,781

FIGURA 50 – PASSO 1 DO DESAFIO 10 (AMOSTRAGEM EM DOIS ESTÁGIOS)  
FONTE: O AUTOR (2014)

#### Passo 2

Da mesma forma que nos desafios do grupo 2, neste conjunto de desafios o aluno deve calcular as estimativas para o inventário piloto no segundo passo, acrescentando as estimativas por unidade primária, conforme ilustrado na Figura 51.

Desafios							« Anterior Próximo »	
Unidade Primária	Média (m <sup>3</sup> / 0,25 ha)	Coef. Variação (%)	Variação (m <sup>3</sup> / 0,25 ha)	Número Unidades	Desvio Padrão	Volume Total		
1	68,105	36,747	626,318	7	25,026	476,733		
2	66,133	79,064	2.733,985	7	52,288	462,934		
3	51,642	40,77	443,294	7	21,055	361,492		
4	83,409	27,868	540,3	7	23,244	583,86		
5	41,581	46,271	370,176	7	19,24	291,068		
6	48,572	13,787	44,843	7	6,696	340,004		
7	53,498	54,332	844,839	7	29,066	374,483		

 Responder

### Estimativas do inventário piloto

Média por subunidade:	56,752	m <sup>3</sup> / 0,25 ha
Variação entre:	50,214	(m <sup>3</sup> / 0,25 ha) <sup>2</sup>
Variação dentro:	745,206	(m <sup>3</sup> / 0,25 ha) <sup>2</sup>
Estimativa da variação total:	795,42	(m <sup>3</sup> / 0,25 ha) <sup>2</sup>
Intensidade de amostragem para unidades secundárias:	3	
Tamanho da população:	4000000	
Intensidade de amostragem para unidades primárias:	21	

 Responder

 Responder todas

ALUNO 40

 Sair

FIGURA 51 – PASSO 2 DO DESAFIO 10 (AMOSTRAGEM EM DOIS ESTÁGIOS)  
 FONTE: O AUTOR (2014)

### Passo 3

No terceiro passo, o sistema mostra ao aluno uma tela semelhante à ilustrada na Figura 50 acrescentando as unidades primárias necessárias ao inventário definitivo. Neste exemplo, o inventário piloto é formado por 18 unidades primárias (Figura 50). A intensidade amostral foi calculada e resultou em 21 unidades (Figura 51). Portanto, o sistema apresentará uma tabela com as 18 unidades primárias anteriores, acrescentando 3 unidades ao final da tabela para formar as 21 unidades do inventário definitivo. O número de unidades secundárias permaneceu o mesmo, pois a intensidade de amostragem calculada foi igual a 3 (Figura 51), o que é menor do que a do inventário piloto, que foi igual a 7 (Figura 50).

### Passo 4

No quarto passo desse grupo de desafios o aluno deve calcular as estimativas para o inventário definitivo. Neste exemplo ele irá calcular as estimativas baseado em uma amostra de 21 unidades primárias, com 7 unidades secundárias. A tela apresentada pelo sistema é semelhante à ilustrada na Figura 51.

#### 4.3.4 Desafios do grupo 4

Nos desafios do grupo 4 são apresentados 2 passos descritos nas seções a seguir.

##### Passo 1

No primeiro passo, uma tabela representando 2 medições em ocasiões diferentes é mostrada ao aluno. Os valores apresentados na tabela estão em  $m^3/ha$ . No exemplo da Figura 52, o sistema sorteia 45 unidades amostrais na primeira ocasião e simula um crescimento para a segunda ocasião, formando, assim, as unidades amostrais permanentes.

**14º - Amostragem em Múltiplas Ocasões - Amostras Com Repetição Total**

Neste desafio a população utilizada foi um povoamento de Eucalyptus sp. Foram feitas duas medições, aqui denominadas de 1ª e 2ª ocasiões. Neste povoamento foram obtidas unidades amostrais aleatórias selecionadas na primeira ocasião. A tabela ao lado ilustra as duas ocasiões e suas respectivas unidades amostrais, totalizando 45 unidades amostrais permanentes. Os valores representados na tabela estão expressos em  $m^3/ha$ .

O objetivo deste desafio é calcular as estimativas para primeira e segunda ocasião, bem como as estimativas para o crescimento, considerando a amostragem com repetição total.

O número de unidades da população na primeira ocasião é de  $N_1=1.500$ .

O número de unidades medidas é de  $n_1=m=45$ . Onde:

$n_1$ =número de unidades amostrais tomadas na primeira ocasião

Unidade Nº	1ª Ocasão	2ª Ocasão
1	127,937	175,697
2	140,233	199,024
3	137,725	194,174
4	127,251	162,017
5	171,985	263,699
6	125,941	159,900
7	171,582	262,813
8	106,064	139,013
9	75,370	95,042
10	124,655	169,661
11	178,825	278,907
12	91,867	107,280
13	182,239	286,632
14	182,305	286,783
15	135,232	189,400
16	74,396	93,695
17	93,487	109,659
18	162,998	244,266
19	96,726	125,312
20	126,733	173,472

Pontuação ALUNO 40 Sair

FIGURA 52 – PASSO 1 DO DESAFIO 14 (AMOSTRAGEM EM MÚLTIPLAS OCASIÕES – AMOSTRAGEM COM REPETIÇÃO TOTAL)  
FONTE: O AUTOR (2014)

##### Passo 2

No segundo passo o aluno deve calcular as estimativas<sup>1</sup> para a primeira e segunda ocasiões, bem como para o crescimento e mudança. A Figura 53 ilustra a tela com os respectivos campos que devem ser respondidos. Nessa figura a segunda ocasião foi suprimida por possuir os mesmos campos da primeira ocasião.

FIGURA 53 – PASSO 2 DO DESAFIO 14 (AMOSTRAGEM EM MÚLTIPLAS OCASIÕES – AMOSTRAGEM COM REPETIÇÃO TOTAL)  
FONTE: O AUTOR (2014)

#### 4.3.5 Sistema de Pontuação

Foi desenvolvido um sistema de pontuação no software para que os alunos pudessem acompanhar o seu desempenho na resolução dos exercícios. Também é possível visualizar uma tabela comparativa entre o desempenho dos alunos. A Figura 54 ilustra um exemplo da tela comparativa da pontuação dos alunos, onde o “ALUNO 30” está conectado no sistema e pode verificar que ele está com 25 pontos, o “ALUNO 20” está com 40 pontos, e o “ALUNO 40” está com 28 pontos.

ALUNO	Pontuação
ALUNO 20	40
ALUNO 40	28
ALUNO 30	25
ALUNO 1	0
ALUNO 2	0

FIGURA 54 – TELA DE TABELA COMPARATIVA DA PONTUAÇÃO  
FONTE: O AUTOR (2014)

A pontuação é somada por acerto, ou seja, a cada campo calculado corretamente na primeira tentativa, são computados 10 pontos. A cada tentativa um ponto é descontado. Por exemplo, se o aluno acertar o valor calculado na terceira tentativa, serão computados 8 pontos. O aluno recebe no mínimo 1 ponto por cálculo correto, mesmo que tenha ultrapassado 9 tentativas.

A pontuação obtida é vista no canto superior direito de cada desafio. No exemplo da Figura 43 a pontuação foi de 50 pontos para o desafio “Parâmetros da População”. Abaixo do nome de cada desafio aparecem as estrelas obtidas em cada um. As estrelas se referem a uma pontuação relativa. No exemplo da Figura 43 o aluno obteve 5 estrelas no desafio 1, pois o seu percentual de acerto ficou entre 80 e 100% das questões do desafio.

#### 4.4 AVALIAÇÃO DO SOFTWARE

A avaliação do *software* foi realizada com uma turma do quinto semestre do curso de graduação de Engenharia Florestal da Universidade Federal do Paraná. O instrumento de avaliação ilustrado na Tabela 2 foi respondido por 23 alunos após o uso do *software* no laboratório de informática.

A Tabela 2 mostra os resultados da avaliação por meio da média e da variância das respostas obtidas em cada questão objetiva.

Os itens 1, 2, 3 e 6 procuraram avaliar a percepção dos alunos quanto ao benefício educacional que o *software* proporciona. Pôde-se observar, pela média obtida nas respostas, que os alunos concordaram com os benefícios apresentados.

A facilidade de uso do *software* foi avaliada pelos itens 4 e 8. Neste grupo os alunos concordaram com o fato de que o *software* é de fácil operação e que os auxilia na solução dos problemas.

TABELA 2 – RESULTADOS DA AVALIAÇÃO (N=23)

No.	Itens Avaliados	Média	DP
1	O Simulador Florestal agrega valor à disciplina de Inventário Florestal.	4,96	0,21
2	O Simulador Florestal possibilita outra estratégia de aprendizado.	4,74	0,54
3	Eu sinto que o Simulador Florestal pode beneficiar o entendimento sobre os cálculos estatísticos referentes aos processos de amostragem.	4,91	0,29
4	O <i>software</i> Simulador Florestal é de fácil operação.	4,78	0,52
5	Outros exercícios ao longo do curso poderiam ser apresentados por um objeto de aprendizagem semelhante a este.	4,48	0,67
6	De um modo geral o Simulador Florestal é um bom complemento ao conteúdo ministrado em sala de aula.	4,78	0,42
7	O sistema de pontuação do <i>software</i> me incentivou a pensar melhor na solução dos problemas propostos.	4,00	1,21
8	A correção dos exercícios em tempo real me ajudou na solução dos problemas.	4,96	0,21
9	Eu estou interessado em utilizar o Simulador Florestal novamente.	4,87	0,46
10	De um modo geral, como você avalia o Simulador Florestal. (Péssimo, Fraco, Regular, Bom e Excelente)	4,61	0,50

FONTE: O AUTOR (2014)

O item 7, que avalia o sistema de pontuação (ver seção 4.3.5), ficou com a menor média entre os quesitos avaliados. Sua tendência foi para uma concordância parcial e gerou a maior discrepância entre as avaliações dos alunos, que pode ser avaliada pelo desvio padrão das respostas. Uma hipótese é o fato de o sistema fornecer um relatório comparando o desempenho dos alunos. Outra hipótese é o fato do sistema penalizar o aluno que insere muitas vezes valores errados no sistema.

Os itens 5 e 9 avaliaram o nível de interesse dos alunos para utilizar este ou outros recursos parecidos no futuro. Neste grupo as respostas revelaram que eles têm interesse neste tipo de abordagem como complemento à aprendizagem.

O item 10 revelou que o *software* avaliado teve uma percepção positiva entre os alunos.

As respostas dos itens abertos podem ser encontradas no Apêndice 8 e confirmaram a percepção positiva entre os estudantes. Os alunos 5 e 8 responderam que não gostaram do sistema de pontuação. Por outro lado os alunos 1 e 20 apontaram o sistema de pontuação como ponto positivo do *software*. A opinião controvertida sobre esse item se refletiu no item 7 do instrumento de avaliação com o maior desvio padrão entre os itens avaliados.

A correção automática foi citada por oito alunos (8, 10, 11, 12, 16, 19, 21 e 22) como ponto positivo do software nos itens abertos (Apêndice 8). Outro ponto positivo apontado pelos alunos foi a facilidade de uso citado por quatro alunos (4, 5, 19 e 20).

Como pontos negativos, nos itens abertos (Apêndice 8), além do sistema de pontuação já citado anteriormente, um aluno colocou a interface, um aluno sugeriu modificar a navegabilidade para facilitar a leitura dos enunciados e outro aluno afirmou que o software não está adequado à realidade dos inventários florestais.

O trabalho apresentado seguiu uma metodologia de avaliação semelhante ao estudo de Nugent *et al.* (2006). Esses autores avaliaram um *software* educacional focado em tópicos de orientação a objetos da ciência da computação. Fizeram a avaliação com 33 alunos, que responderam a um instrumento de avaliação com 9 perguntas, e cinco níveis de resposta para cada pergunta (Discordo Totalmente, Discordo, Neutro, Concordo e Concordo Totalmente). Eles utilizaram o mesmo método aplicado neste trabalho, ou seja, atribuíram a cada resposta um valor numérico (de 1 a 5, respectivamente) e tabularam as respostas obtendo a média e o desvio padrão. Avaliaram aspectos como: facilidade de uso do *software*, benefícios educacionais, interesse futuro e avaliação geral. O resultado obtido por esses autores foi semelhante ao resultado obtido neste trabalho, pois relataram que: os estudantes acharam o *software* simples para ser utilizado, agregando valor ao curso e mantendo-os interessados. Outro aspecto interessante que pode ser comparado foi que a avaliação geral do *software* desenvolvido e testado por Nugent *et al.* (2006) obteve uma média de 4,31 com desvio padrão de 0,53, enquanto que o *software* apresentado neste trabalho obteve uma média de 4,61 com desvio padrão de 0,50.

O trabalho de Romero-Troncoso *et al.* (2014) também possui uma metodologia de avaliação semelhante a este. Nesse trabalho eles desenvolveram e avaliaram uma ferramenta para o estudo e análise de alguns aspectos de motores de indução. Segundo os autores, um grupo de 18 estudantes do curso de graduação de engenharia elétrica participaram da avaliação de um *software* educacional, a qual



contou com 5 questões objetivas. Romero-Troncoso *et al.* (2014) utilizaram o mesmo método de avaliação formalizado por Nugent *et al.* (2006) e por este trabalho, com cinco níveis de respostas e um valor numérico de 1 a 5 atribuído a cada resposta. Romero-Troncoso *et al.* (2014) concluíram que os resultados sugeridos na avaliação produziram um impacto positivo nos estudantes. Segundo esses autores a nota média obtida na avaliação foi maior do que 4,0 em uma escala de 5,0, indicando uma experiência positiva no aprendizado e interesse no assunto estudado.

## 5 CONCLUSÕES

Este trabalho apresentou o desenvolvimento e a avaliação de um *software* de aprendizagem para os processos de amostragem da disciplina de Inventário Florestal. A aplicação é acessada pela Internet, exercita 16 processos de amostragem com os estudantes, faz a correção dos exercícios de forma *on-line* e propõe dados para cada estudante nos enunciados.

Os resultados da avaliação do *software* demonstraram que a percepção dos alunos em relação ao benefício educacional, à facilidade de uso e ao interesse futuro foi positiva e que o *software* pode ser aplicado nas disciplinas de inventário florestal como complemento ao conteúdo. A avaliação indicou também que o sistema de pontuação e comparação de desempenho teve concordância parcial entre os alunos.

O uso de tecnologia da informação e comunicação nas disciplinas dos cursos de Engenharia, em especial na Engenharia Florestal, é muito limitado ou inexistente, de acordo com as pesquisas realizadas neste trabalho. Por outro lado, tem um potencial bastante grande, conforme demonstrado pelos resultados obtidos.

Recomenda-se o uso do *software* desenvolvido e apresentado neste trabalho como parte integrante no programa da disciplina de Inventário Florestal de cursos de graduação ou pós-graduação.

Uma biblioteca de programação para processos de amostragem em inventário florestal foi projetada e construída com base na modelagem orientada a objetos. Essa biblioteca foi construída na linguagem de programação Java e é composta por 31 classes que podem ser utilizadas para cálculos das estimativas de um inventário florestal, por meio de qualquer processo de amostragem. Também pode ser utilizada para a implementação de um *software* específico de inventário florestal que disponibilize os processos de amostragens para seus usuários.

O método “núcleo estimador variável” foi utilizado para geração de dados pseudoaleatórios, possibilitando que cada estudante tivesse seu próprio conjunto de

dados para realizar os exercícios propostos pelo *software*. Foi realizado um estudo desse método, que comprovou ser possível estimar funções densidade de probabilidade a partir de conjuntos de dados que tenham características não aderentes a uma função conhecida. Portanto, pode-se aplicar esse método a dados que tenham características bi ou multimodais para se encontrar e estudar a fdp gerada por esses dados.

Finalmente, pode-se afirmar que este trabalho cumpriu o objetivo geral de desenvolver e avaliar um *software* educacional para a disciplina de Inventário Florestal, bem como os objetivos específicos: implementar e testar uma biblioteca de programação orientada a objetos, na linguagem Java, que realize os cálculos estatísticos de processos de amostragem; construir e testar um gerador de dados capaz de produzir dados aleatórios, os quais respeitem uma determinada distribuição, para que cada aluno tenha seu próprio conjunto de dados nos exercícios propostos pelo sistema; disponibilizar a aplicação gerada na internet, para que os alunos tenham acesso ao *software* por meio de um navegador e avaliar o *software* gerado com uma turma da graduação em Engenharia Florestal da Universidade Federal do Paraná.

## 6 RECOMENDAÇÕES PARA TRABALHOS FUTUROS

Com base no trabalho apresentado pode-se sugerir como trabalhos futuros:

- Criação de um módulo do *software* educacional que complete o *software* apresentado, incluindo os métodos de amostragens e integrando-os aos processos de amostragens;
- Ampliar a avaliação do software para outros aspectos, tais como usabilidade (IHC – Interface Humano-Computador), navegabilidade, apresentação das informações e outros aspectos pedagógicos;
- Avaliar o software gerado com mais turmas da disciplina de Inventário Florestal, incluindo cursos de Engenharia Florestal de outras instituições de ensino;
- Ampliar as funcionalidades do software para: incluir material instrucional ou links sobre os processos de amostragem, com exemplos, figuras, imagens e vídeos; e criar um help-online;
- Modelagem e implementação de uma biblioteca de programação para os métodos de amostragem em inventário florestal;
- Utilização da biblioteca de programação criada na avaliação de inventários florestais reais;
- Utilização do método “núcleo estimador variável” para estimar funções densidade de probabilidade de dados biométricos florestais;
- Estudo da utilização do método “núcleo estimador variável” para simulação de dados biométricos florestais, como, por exemplo, a simulação de crescimento de uma floresta com características bimodais.

## REFERÊNCIAS

ALEXANDRE, B. T. **Introduction to Nonparametric Estimation**. New York: Springer-Verlag, 2009.

ALTMAN, N.; LÉGER, C. Bandwidth selection for kernel distribution function estimation. **Journal of Statistical Planning and Inference**, 15 August 1994. 195-214.

BARANZANO, R. H. **Non-parametric kernel density estimation-based permutation test: Implementation and comparisons**. Tese de Doutorado. Uppsala Universitet. Uppsala, Suécia, p. 52. 2011.

BARNES, J. D.; KÖLLING, M. **Programação orientada a objetos com Java**. Tradução de Edsib Furmankiewicz e João Luiz Silva Barbosa. São Paulo: Pearson Prentice Hall, 2009.

BARRETTO, S. F. A.; PIAZZALUNGA, R.; RIBEIRO, V. G. A web-based 2D structural analysis educational software. **Computer Applications in Engineering Education**, v. 11, p. 83-92, 2003.

BECK, K. et al. Manifesto for Agile Software Development, 2001. Disponível em: <<http://agilemanifesto.org/>>.

BOEHM, B. A Spiral Model for Software Development and Enhancement. **Computer**, 21, n. 5, 1988. 61-72.

BOOCH, G.; RUMBAUGH, J.; JACOBSON, I. **UML - Guia do Usuário**. Tradução de Fábio Freitas da Silva e Cristina de Amorim. Rio de Janeiro: Elsevier, 2012.

BOURQUE, P.; FAIRLEY, R. E. **Guide to the Software Engineering Body of Knowledge, Version 3.0**. IEEE Computer Society. Los Alamitos. 2014.

BRASIL. Resolução n. 3, de 2 de fevereiro de 2006. Institui as Diretrizes Curriculares Nacionais para o curso de graduação em Engenharia Florestal e dá outras providências. **Diário Oficial da União**, Brasília, DF, v. I, p. 33-34, 2006.

BRASIL. Banco Internacional de Objetos Educacionais, 2008. Disponível em: <<http://objetoseducacionais2.mec.gov.br/>>. Acesso em: 28 fev. 2014.

BRENT, R. P. **Algorithms for Minimization without Derivatives**. Englewood Cliffs N.J: Prentice-Hall, 1973.

CIENTEC. **Software Mata Nativa**, 2014. Disponível em: <<http://www.matanativa.com.br/>>. Acesso em: 10 mar. 2014.

COCHRAN, W. G. **Sampling Techniques**. third edition. ed. New York: John Wiley & Sons, 1977.

DAVIS, D.; IHAKA, R.; FENSTERMACHER, P. **Cryptographic randomness from air turbulence in disk drives**. CRYPTO '94. New York: Springer-Verlag. 1994. p. 114-120.

DESHPANDE, A. A.; HUANG, S. H. Simulation Games in Engineering Education: A State-of-the-Art Review. **Computer Applications in Engineering Education**, Cincinnati, Ohio, 2009. 399-410.

DEVROYE, L. **Non-Uniform Random Variate Generation**. New York: Springer-Verlag, 1986. 843 p.

DJORDJEVIC, J.; NIKOLIC, B.; MILENKOVIC, A. Flexible web-based educational system for teaching computer architecture and organization. **IEEE Transactions on Education**, v. 48, n. 2, p. 264-273, 2005.

DRUSZCZ, J. P. et al. Eficiência de duas variações estruturais do método de amostragem de área fixa em plantações de Pinus taeda. **Revista Floresta**, Curitiba, dez 2013. 621-632.

EZMORTH. **EZMorth**, 2014. Disponível em: <<http://ezmorph.sourceforge.net/index.html>>. Acesso em: 17 mar. 2014.

FIX, E.; HOGES, J. **Nonparametric discrimination: consistency properties**. Report Number4, USAF School of Aviation Medicine. Texas. 1951.

FIX, E.; HOGES, J. Discriminatory analysis. Nonparametric discrimination: consistency properties. **International Statistical Review/Revue Internationale de Statistique**, v. 57, n. 3, p. 238-247, 1989.

FOREST, E. The ADDIE Model: Instructional Design. **Educational Technology**, 2014. Disponível em: <<http://educationaltechnology.net/the-addie-model-instructional-design/>>. Acesso em: 19 mar. 2014.

GENTLE, J. E. **Random Number Generation and Monte Carlo Methods**. 2a. ed. New York: Springer, 2003.

GILB, T.; FINZI, S. **Principles of Software Engineering Management**. Reading, MA: Addison-Wesley, 1988.

GUIDOUM, A. C. Kernel Estimator and Bandwidth Selection for Density. **The Comprehensive R Archive Network**, 14 Agosto 2013. Disponível em: <<http://cran.r-project.org/web/packages/kedd/vignettes/kedd.pdf>>. Acesso em: 27 mar. 2014.

HAAR, M. Introduction to Randomness and Random Numbers. **RANDOM.ORG**, 2014. Disponível em: <<http://www.random.org/randomness/>>. Acesso em: 26 julho 2014.

HANSMANN, R. et al. Enhancing environmental awareness: Ecological and economic effects of food consumption. **Simulation & Gaming**, v. 36, n. 3, p. 364-382, 2005.

HIROSE, Y.; SUGIURA, J.; SHIMOMOTO, K. Industrial waste-management simulation game and its educational effects. **J Material Cycles Waste Manag**, v. 6, p. 58-63, 2004.

HUSH, B.; MILLER, C. I.; BEERS, T. W. **Forest Mensuration**. 3rd.. ed. New York: J. Wiley, 1982. 337 p.

IEEE. IEEE Standards Collection: Software Engineering. **IEEE Standard 610.12-1990**, 1993.

INSTITUTO BRASILEIRO DE FLORESTAS. **Software Faça Floresta**, 2014. Disponível em: <<http://www.facafloresta.com.br/index.php>>. Acesso em: 10 mar. 2014.

JACOBSON, I.; BOOCH, G.; RUMBAUGH, J. **The Unified Software Development Process**. Reading: Addison Wesley, 1999.

JANDL JUNIOR, P. **Java: Guia do Programador**. São Paulo: Novatec, 2007.

JEFFERY, S. S. **Smoothing Methods in Statistics**. New York: Springer-Verlag, 1996.

JEHNG-JUNG, K.; YING-JIE, C. An educational water qualitymanagement game. **Water Sci Technol**, v. 34, n. 12, p. 205-211, 1996.

KAY, R. H.; KNAACK, L. Evaluating the learning in learning objects. **Open Learning**, v. 22, n. 1, p. 5-28, Fevereiro 2007.

KÖHL, M.; MAGNUSSEN, S. S.; MARCHETTI, M. **Sampling Methods, Remote Sensing and GIS Multiresource Forest Inventory**. Berlin: Springer-Verlag, 2006. 373 p.

LAAR, A. V.; AKÇA, A. **Forest Mensuration**. 2nd. ed. Gottenberg: Springer, v. 13, 2007. 383 p.

LASDON, L. S. et al. Design and testing of a generalized reduced gradient code for nonlinear programming. **ACM Transactions on Mathematical Software (TOMS)**, v. 4, n. 1, p. 34-50, 1978.

LIMA, A. D. S. **UML 2.3: do requisito à solução**. São Paulo: Érica, 2011.

MACHADO, S. D. A.; FIGUEIREDO FILHO, A. **Dendrometria**. Curitiba: UFPR, 2003.

MATSUMOTO, M.; NISHIMURA, T. Mersenne Twister: A 623-dimensionally equidistributed uniform pseudo-random number generator. **ACM Transactions on Modeling and Computer Simulation**, n. 8, 1998. 3-30.

MUÑOZ-ABELLA, B.; ÁLVAREZ-CALDAS, C.; RUBIO, L. Computer-aided tool for teaching mechanical clutch systems design. **Computer Applications in Engineering Education**, v. 19, p. 493-500, 2011.

NADARAYA, E. A. On estimation regression. **Theory of Probability & Its Applications**, 3, 1964. 497-500.

NERUR, S.; MAHAPATRA, R.; MANGALARAJ, G. Challenges of migrating to agile methodologies. **Communications of the ACM**, 48, n. 5, 2005. 72-78.

NUGENT, G.; SOH, L.-K.; SAMAL, A. Desing, development, and validation of learning objects. **Journal of Educational Technology Systems**, v. 34, n. 3, p. 271-281, 2006.

ORACLE. **The Java Tutorials**, 2014. Disponível em: <<http://docs.oracle.com/javase/tutorial/getStarted/intro/index.html>>. Acesso em: 17 mar. 2014.

ORACLE. **MySQL**, 2014. Disponível em: <<http://dev.mysql.com/downloads/mysql/>>. Acesso em: 17 mar. 2014.

PARZEN, E. On estimation of a probability density function and mode. **The annals of mathematical statistics**, v. 33, p. 1075-1076, 1962.

PAUL, R. J.; BALMER, D. W. **Simulation Modelling**. Lund: Chartwell-Bratt Student-Text Series, 1993.

PÉLLICO NETTO, S.; BRENA, D. A. **Inventário Florestal**. Curitiba: Universidade Federal do Paraná, 1997. 316 p.

POPPENDIECK, M.; CUSUMANO, M. A. Lean software development: A tutorial. **Software, IEEE**, v. 29, n. 5, p. 26-32, 2012.

PRESSMAN, R. S. **Engenharia de Software**. Sexta Edição. ed. São Paulo: McGraw-Hill, 2006.

PRODAN, M. et al. **Mensura Florestal**. Tradução de Claudia Eppelin. San José: Imprensa del Instituto Interamericano de Cooperación para la Agricultura, 1997. 586 p.

PROJETO CR CAMPEIRO - UFSM. **C7 LDFP**, 2014. Disponível em: <[http://www.crcampeiro.net/novo/Pages/apps\\_android](http://www.crcampeiro.net/novo/Pages/apps_android)>. Acesso em: 10 mar. 2014.



QUEIROZ, W. T. D. **Amostragem em inventário florestal**. Belém: Editora Universidade Federal Rural da Amazônia, 2012. 441 p.

R CORE TEAM. R: A language and environment for statistical. **R Foundation for Statistical Computing, Viena, Austria**, Vienna, Austria, 2014. Disponível em: <<http://www.r-project.org/>>.

ROMERO-TRONCOSO, R. J. et al. Experimental system for teaching induction motor faults during the startup transient and steady state. **Computer Applications in Engineering Education**, v. 22, p. 33-38, 2014.

ROSELLÓ, E. G. et al. Visual Wavelet-Lab: An object-oriented library and a GUI application for the study of the wavelet transform. **Computer Applications in Engineering Education**, v. 22, p. 23-32, 2014.

ROSENBATT, M. Remarks on some nonparametric estimates of a density function. **The Annals of Mathematical Statistics**, v. 27, p. 832-837, 1956.

SAIN, S. R.; SCOTT, D. W. On locally adaptive density estimation. **Journal of the American Statistical Association**, v. 91, n. 436, p. 1525-1534, 1996.

SANQUETTA, C. R. et al. **Inventários Florestais: Planejamento e Execução**. 2a. ed. Curitiba: Multi-Graphic Gráfica e Editora, 2009.

SANTOS, F.; PÉREZ, M. F. M.; MANTEIGA, W. G. Forest fire spatial pattern analysis in Galicia (NW Spain). **Journal of Environmental Management**, 128, 2013. 30-42.

SCOTT, D. W. **Multivariate Density Estimation: Theory, Practice, and Visualization**. New York: John Wiley, 1992.

SCOTT, D. W. **Multivariate Density Estimation and Visualization**. Humboldt-Universität Berlin. Berlin, p. 27. 2004.

SCOTT, D.; TERRELL, G. Biased and unbiased cross-validation in density estimation. **Journal of the American Statistical Association**, 82, 1987. 1131-1146.

SENCHA. **Ext JS 3.4.0**, 2014. Disponível em: <<http://www.sencha.com/products/extjs3>>. Acesso em: 17 mar. 2014.

SHEATHER, S. J.; JONES, M. C. A Reliable Data-Based Bandwidth Selection Method for Kernel Density Estimation. **Journal of the Royal Statistical Society. Series B (Methodological)**, v. 53, p. 683-690, 1991.

SHIUE, C. J. Systematic sampling with multiple random starts. **Forest Science**, v. 6, n. 1, p. 42-50, 1960.

SILVERMAN, B. W. **Density Estimation for Statistics and Data Analysis**. London: Chapman & Hall., 1986.

SOUZA, A. R. D.; YONEZAWA, W. M.; SILVA, P. M. D. Desenvolvimento de habilidades em tecnologia da informação e comunicação (TIC) por meio de objetos de aprendizagem. In: PRATA, C. L.; NASCIMENTO, A. C. A. D. A. **Objetos de Aprendizagem**. Brasília: SEED, 2007. p. 154.

STURGES, H. A. The choice of a class interval. **Journal of the American Statistical Association**, v. 21, n. 153, p. 65-66, 1926.

TERRELL, G. R.; SCOTT, D. W. Variable Kernel Density Estimation. **The Annals of Statistics**, 20, n. 3, Sep. 1992. 1236-1265.

THE APACHE SOFTWARE FOUNDATION. Apache Commons. **Apache Commons**, 2013. Disponível em: <<http://commons.apache.org/>>. Acesso em: 17 mar. 2014.

THE APACHE SOFTWARE FOUNDATION. Apache Commons Math, Release 3.3. **Apache Commons Math**, 2013. Disponível em: <<http://commons.apache.org/proper/commons-math/index.html>>. Acesso em: 17 mar. 2014.

THE APACHE SOFTWARE FOUNDATION. Apache Tomcat. **Apache Tomcat**, 2014. Disponível em: <<http://tomcat.apache.org/>>. Acesso em: 2014.

THE APACHE SOFTWARE FOUNDATION. The Apache POI Project. **Apache POI - the JAVA API for Microsoft Documents**, 2014. Disponível em: <<http://poi.apache.org/>>. Acesso em: 17 mar. 2014.

TIOBE SOFTWARE. **TIOBE Index for March 2014**, 2014. Disponível em: <<http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html>>. Acesso em: 17 mar. 2014.

TORRES, M.; MACEDO, J. Learning sustainable development with a new simulation game. *Simulation & Gaming*, v. 31, n. 1, p. 119-126, 2000.

TREESOFTWARE. **TreeSoftware**, 2013. Disponível em: <<http://treesoftware.lwsite.com.br/>>. Acesso em: 10 ago. 2014.

VON NEUMANN, J. Various techniques used in connection with random digits. **Applied Math Series**, 1, n. 12, 1951. 36-38.

WAND, M. P.; JONES, M. C. **Kernel Smoothing**. London: Chapman and Hall, 1995.

WANG, Z. et al. A virtual environment simulator for mechanical system dynamics with online interactive control. **Advances in Engineering Software**, v. 37, n. 10, p. 631-642, 2006.

WARE, K. D.; CUNIA, T. Continuous forest inventory with partial replacement of samples. **Forest science**, Monograph 3, 1962. 40p.

WICHMANN, B. A.; HILL, I. D. Algorithm AS 183: An Efficient and Portable Pseudo-random Number Generator. **Applied Statistics**, n. 31, 1982. 188-190.

WILEY, D. A. Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy. **The Instructional Use of Learning Objects**: Online Version., 2000. Disponível em: <<http://reusability.org/read/chapters/wiley.doc>>. Acesso em: 27 fev. 2014.

WOLFGANG, H. et al. **Nonparametric and Semiparametric**. Berlin Heidelberg: Springer-Verlag, 2004.

## APÊNDICE 1 – ESTUDO DOS MOMENTOS DA FUNÇÃO NÚCLEO ESTIMADOR VARIÁVEL

As mudanças de variáveis apresentadas em (19) e (20) são utilizadas no estudo dos momentos da função núcleo estimador variável.

$$u = \frac{(x - X_i)}{h\lambda_i} \quad \text{ou} \quad (19)$$

$$x = X_i + uh\lambda_i$$

$$\frac{dx}{du} = h\lambda_i \therefore dx = h\lambda_i du \quad (20)$$

Considerando-se que as mudanças de variáveis propostas nas equações em (16) e (17), e que a função núcleo estimador variável é gaussiana, conforme apresentada na equação (12) da seção 2.3.2, pode-se reescrever a função núcleo definida em (12) como demonstrado em (18).

$$K(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} \quad (21)$$

Isso torna a equação (18) uma função gaussiana normal de média 0 e variância 1 em relação a  $u$ . Portanto, podem-se assumir os momentos de  $K(u)$ , conforme apresentado em (19), (20) e (21).

$$E\{u^0\} = \int_{-\infty}^{+\infty} K(u) du = 1 \quad (22)$$

$$E\{u\} = \int_{-\infty}^{+\infty} u \cdot K(u) du = 0 \quad (23)$$

$$E\{u^2\} = \int_{-\infty}^{+\infty} u^2 \cdot K(u) du = 1 \quad (24)$$

Considerando as mudanças de variáveis em (19) e (20), e os momentos da função  $K(u)$  apresentados em (22), (23) e (24) pode-se calcular numericamente os momentos da função densidade  $\hat{f}_s(x)$ , conforme apresentado a seguir.

Em (25) está demonstrado o cálculo do primeiro momento da função núcleo estimador variável  $\hat{f}_s(x)$ .

$$\begin{aligned}
\bar{X}_K &= E\{X\} = \int_{-\infty}^{+\infty} x \cdot \hat{f}_s(x) dx = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} x \cdot \frac{1}{h\lambda_i} \cdot K\left(\frac{x - X_i}{h\lambda_i}\right) dx \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i + uh\lambda_i) \cdot \frac{1}{h\lambda_i} K(u) h\lambda_i du = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i + uh\lambda_i) \cdot K(u) du \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i \cdot K(u) + uh\lambda_i \cdot K(u)) du = \frac{1}{n} \sum_{i=1}^n \left[ \int_{-\infty}^{+\infty} X_i \cdot K(u) du + \int_{-\infty}^{+\infty} uh\lambda_i \cdot K(u) du \right] = \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} X_i \cdot K(u) du + \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} uh\lambda_i \cdot K(u) du \\
&= \frac{1}{n} \sum_{i=1}^n X_i \underbrace{\int_{-\infty}^{+\infty} K(u) du}_1 + \frac{1}{n} \sum_{i=1}^n h\lambda_i \underbrace{\int_{-\infty}^{+\infty} u \cdot K(u) du}_0 = \frac{1}{n} \sum_{i=1}^n X_i \therefore \\
E\{X\} &= \frac{1}{n} \sum_{i=1}^n X_i
\end{aligned} \tag{25}$$

Conforme demonstrado, pode-se afirmar que a média da densidade estimada por (22) é a média da amostra de  $X_i$ , apresentada na equação (26).

$$\bar{X}_K = \frac{1}{n} \sum_{i=1}^n X_i \tag{26}$$

Em (27) está demonstrado o cálculo do segundo momento da função núcleo estimador variável  $\hat{f}_s(x)$ .

$$\begin{aligned}
E\{X^2\} &= \int_{-\infty}^{+\infty} x^2 \cdot \hat{f}_s(x) dx = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} x^2 \cdot \frac{1}{h\lambda_i} \cdot K\left(\frac{x - X_i}{h\lambda_i}\right) dx \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i + uh\lambda_i)^2 \cdot K(u) du = \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i^2 + 2X_iuh\lambda_i + u^2h^2\lambda_i^2) \cdot K(u) du \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} (X_i^2 K(u) + 2X_iuh\lambda_i K(u) + u^2h^2\lambda_i^2 K(u)) du \\
&= \frac{1}{n} \sum_{i=1}^n \left[ \int_{-\infty}^{+\infty} X_i^2 K(u) du + \int_{-\infty}^{+\infty} 2X_iuh\lambda_i K(u) du + \int_{-\infty}^{+\infty} u^2h^2\lambda_i^2 K(u) du \right] \\
&= \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} X_i^2 K(u) du + \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} 2X_iuh\lambda_i K(u) du + \frac{1}{n} \sum_{i=1}^n \int_{-\infty}^{+\infty} u^2h^2\lambda_i^2 K(u) du \\
&= \frac{1}{n} \sum_{i=1}^n X_i^2 \underbrace{\int_{-\infty}^{+\infty} K(u) du}_1 + \frac{2}{n} \sum_{i=1}^n X_i h\lambda_i \underbrace{\int_{-\infty}^{+\infty} u K(u) du}_0 + \frac{1}{n} \sum_{i=1}^n h^2\lambda_i^2 \underbrace{\int_{-\infty}^{+\infty} u^2 K(u) du}_1 \therefore \\
E\{X^2\} &= \frac{1}{n} \sum_{i=1}^n X_i^2 + \frac{1}{n} \sum_{i=1}^n h^2\lambda_i^2
\end{aligned} \tag{27}$$

A variância da densidade  $\hat{f}_s(x)$  é calculada conforme demonstrado em (28).

$$\begin{aligned}
S_K^2 &= E\{X^2\} - (E\{X\})^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 + \frac{1}{n} \sum_{i=1}^n h^2 \lambda_i^2 - \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^2 \\
&= \underbrace{\frac{1}{n} \sum_{i=1}^n X_i^2 - \left( \frac{1}{n} \sum_{i=1}^n X_i \right)^2}_{\hat{\sigma}^2} + \frac{1}{n} \sum_{i=1}^n h^2 \lambda_i^2 \therefore \\
S_K^2 &= \hat{\sigma}^2 + \frac{1}{n} \sum_{i=1}^n h^2 \lambda_i^2
\end{aligned} \tag{28}$$

Para obter a variância estimada da amostra, procede-se à transformação apresentada na equação (29).

$$\frac{n}{n-1} \hat{\sigma}^2 = \hat{s}^2 \tag{29}$$

Aplicando-se a transformação apresentada em (29) na equação obtida em (28) tem-se a variância estimada da amostra, conforme apresentado em (30).

$$\begin{aligned}
s_K^2 &= \hat{\sigma}^2 \cdot \left( \frac{n}{n-1} \right) + \frac{1}{n} \sum_{i=1}^n h^2 \lambda_i^2 \therefore \\
s_K^2 &= \hat{s}^2 + \frac{1}{n} \sum_{i=1}^n h^2 \lambda_i^2
\end{aligned} \tag{30}$$

## APÊNDICE 2 – EXEMPLO DE USO DO MÉTODO NÚCLEO ESTIMADOR VARIÁVEL

Para exemplificar o uso do núcleo estimador de densidade variável será tomado o diâmetro de 10 árvores de uma espécie qualquer: 27,0; 27,5; 28,9; 29,8; 30,4; 34,4; 37,4; 38,5; 39,8 e 44,9 e configurado o histograma dessas unidades amostrais. Para definir a quantidade de classes do histograma utilizou-se a regra de Sturges (1926). A Figura 55, por meio das linhas pontilhadas, ilustra o histograma das amostras exemplificadas.

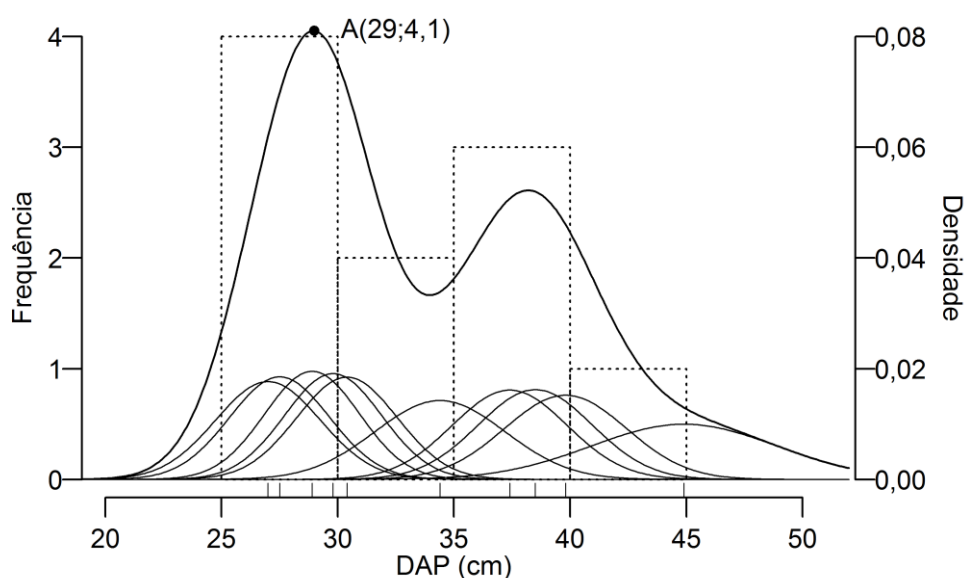


FIGURA 55 – EXEMPLO DO USO DE NÚCLEO ESTIMADORES VARIÁVEL  
FONTE: O AUTOR (2014)

O primeiro passo é calcular as estimativas para a função piloto do núcleo estimador. A função piloto é definida pelo núcleo estimador fixo. O parâmetro de suavidade  $h$  foi obtido segundo o método de Sheather e Jones (1991), por meio do *software* estatístico R (R CORE TEAM, 2014). A função  $\tilde{f}(X_i)$  é definida considerando  $h=2,455$  e  $n=10$ . Substituindo estes valores em (5), pode-se afirmar que a função densidade piloto é descrita pela equação em (31).

$$\begin{aligned}\tilde{f}(x) &= \frac{1}{10 \cdot 2,455} \sum_{i=1}^{10} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-X_i)^2}{2(2,455)^2}} \therefore \\ \tilde{f}(x) &= 0,016 \sum_{i=1}^{10} e^{-\frac{(x-X_i)^2}{12,01}}\end{aligned}\quad (31)$$

Para calcular  $\tilde{f}(X_1) = \tilde{f}(27)$ , procede-se:

$$\begin{aligned}\tilde{f}(27) &= 0,016 \cdot \left( e^{-\frac{(27-27)^2}{12,01}} + e^{-\frac{(27-27,5)^2}{12,01}} + e^{-\frac{(27-28,9)^2}{12,01}} + e^{-\frac{(27-29,8)^2}{12,01}} + e^{-\frac{(27-30,4)^2}{12,01}} \right. \\ &\quad \left. + e^{-\frac{(27-34,4)^2}{12,01}} + e^{-\frac{(27-37,4)^2}{12,01}} + e^{-\frac{(27-38,5)^2}{12,01}} + e^{-\frac{(27-39,8)^2}{12,01}} + e^{-\frac{(27-44,9)^2}{12,01}} \right) = 0,06\end{aligned}\quad (32)$$

Considerando que  $g$  é a média geométrica de  $\tilde{f}(X_i)$  temos, pela equação 8, que  $g=0,05$ . O valor padrão de  $\alpha$  proposto por Silverman (1986) é 0,5. Portanto, para se calcular o fator de localidade  $\lambda_1$  procede-se conforme a equação 33.

$$\lambda_1 = \left( \frac{\tilde{f}(X_1)}{g} \right)^{-\alpha} = \left( \frac{0,06}{0,05} \right)^{-0,5} = 0,92 \quad (33)$$

O valor da janela  $h$  é o mesmo obtido para o núcleo estimador piloto  $\tilde{f}(x)$ , portanto  $h=2,455$  e  $n=10$ . A equação (34) ilustra o cálculo para  $x=29$ cm, com base nos resultados obtidos na Tabela 3.

$$\begin{aligned}\hat{f}(29) &= 0,016 \sum_{i=1}^{10} \frac{1}{\lambda_i} e^{-\frac{(29-X_i)^2}{12,44 \lambda_i^2}} \therefore \\ \hat{f}(29) &= 0,016 \cdot \left( \frac{1}{0,92} \cdot e^{-\frac{(29-27,0)^2}{12,01 \cdot 0,92^2}} + \frac{1}{0,88} \cdot e^{-\frac{(29-27,5)^2}{12,01 \cdot 0,88^2}} + \frac{1}{0,83} \cdot e^{-\frac{(29-28,9)^2}{12,01 \cdot 0,83^2}} + \frac{1}{0,85} \cdot e^{-\frac{(29-29,8)^2}{12,01 \cdot 0,85^2}} \right. \\ &\quad + \frac{1}{0,88} \cdot e^{-\frac{(29-30,4)^2}{12,01 \cdot 0,88^2}} + \frac{1}{1,14} \cdot e^{-\frac{(29-34,4)^2}{12,01 \cdot 1,14^2}} + \frac{1}{1,01} \cdot e^{-\frac{(29-37,4)^2}{12,01 \cdot 1,01^2}} + \frac{1}{1,00} \cdot e^{-\frac{(29-38,5)^2}{12,01 \cdot 1,00^2}} \\ &\quad \left. + \frac{1}{1,07} \cdot e^{-\frac{(29-39,8)^2}{12,01 \cdot 1,07^2}} + \frac{1}{1,63} \cdot e^{-\frac{(29-44,9)^2}{12,01 \cdot 1,63^2}} \right) = 0,08\end{aligned}\quad (34)$$

Observe-se que o valor da densidade calculada para 29 cm é a soma de todas as densidades das 10 curvas gaussianas para  $x=29$ .



TABELA 3 – EXEMPLO DE CÁLCULO DO NÚCLEO ESTIMADOR VARIÁVEL  
 FONTE: O AUTOR (2014)

$X_i$	$\hat{f}(X_i)$	$\lambda_i$	29
27,0	0,06	0,92	0,01
27,5	0,06	0,88	0,01
28,9	0,07	0,83	0,02
29,8	0,07	0,85	0,02
30,4	0,06	0,88	0,01
34,4	0,04	1,14	0,00
37,4	0,05	1,01	0,00
38,5	0,05	1,00	0,00
39,8	0,04	1,07	0,00
44,9	0,02	1,63	0,00
$\hat{f}(29)$			0,08

Com o objetivo de ilustrar a formação da função densidade de probabilidade, a Figura 55 destaca o ponto A que pertence à curva. O cálculo da densidade para 29 cm de DAP (ponto A) está descrito na equação 34. Para encontrarmos o valor da frequência devemos multiplicar o valor da densidade encontrada pelo tamanho da amostra (10) e pelo tamanho da classe (5), resultando em 4,1, conforme representado no gráfico da Figura 55 para o ponto A.

Qualquer ponto da curva pode ser calculado de maneira análoga ao demonstrado para o ponto A.

Para o cálculo da probabilidade acumulada, devem-se obter as probabilidades acumuladas de cada uma das unidades amostrais. Para exemplificar, calcular-se-á, em (32) e (33), a probabilidade de uma árvore ter entre 30 e 35 cm de diâmetro, segundo a função de probabilidade estimada pelo núcleo estimador variável com as 10 unidades amostrais anteriores, conforme a equação 35 e 36.

$$\begin{aligned}
 P(30 < x \leq 35) &= \hat{F}(35) - \hat{F}(30) \\
 \hat{F}(30) &= \frac{1}{10} \cdot \left( F_{\mu=27,0;\sigma=2,455-0,92}(30) + F_{\mu=27,5;\sigma=2,455-0,88}(30) + F_{\mu=28,9;\sigma=2,455-0,83}(30) \right. \\
 &\quad + F_{\mu=29,8;\sigma=2,455-0,85}(30) + F_{\mu=30,4;\sigma=2,455-0,88}(30) + F_{\mu=34,4;\sigma=2,455-1,14}(30) \\
 &\quad + F_{\mu=37,4;\sigma=2,455-1,01}(30) + F_{\mu=38,5;\sigma=2,455-1,00}(30) + F_{\mu=39,8;\sigma=2,455-1,07}(30) \\
 &\quad \left. + F_{\mu=44,9;\sigma=2,455-1,63}(30) \right) = \\
 &= \frac{1}{10} \cdot (0,91 + 0,88 + 0,70 + 0,54 + 0,43 + 0,06 + 0,00 + 0,00 + 0,00 + 0,00) = 0,35
 \end{aligned} \tag{35}$$

$$\begin{aligned}
\hat{F}(35) &= \frac{1}{10} \cdot (F_{\mu=27,0;\sigma=2,455-0,92}(35) + F_{\mu=27,5;\sigma=2,455-0,88}(35) + F_{\mu=28,9;\sigma=2,455-0,83}(35) \\
&\quad + F_{\mu=29,8;\sigma=2,455-0,85}(35) + F_{\mu=30,4;\sigma=2,455-0,88}(35) + F_{\mu=34,4;\sigma=2,455-1,14}(35) \\
&\quad + F_{\mu=37,4;\sigma=2,455-1,01}(35) + F_{\mu=38,5;\sigma=2,455-1,00}(35) + F_{\mu=39,8;\sigma=2,455-1,07}(35) \\
&\quad + F_{\mu=44,9;\sigma=2,455-1,63}(35)) = \\
&= \frac{1}{10} \cdot (1,00 + 1,00 + 1,00 + 0,99 + 0,98 + 0,58 + 0,17 + 0,08 + 0,03 + 0,01) = 0,58 \\
P(30 < x \leq 35) &= 0,58 - 0,35 = 0,23
\end{aligned} \tag{36}$$

Para este exemplo, a probabilidade calculada de uma árvore ter entre 30 e 35 cm de diâmetro é de 23%. A Figura 56 destaca a área da curva calculada para essa probabilidade.

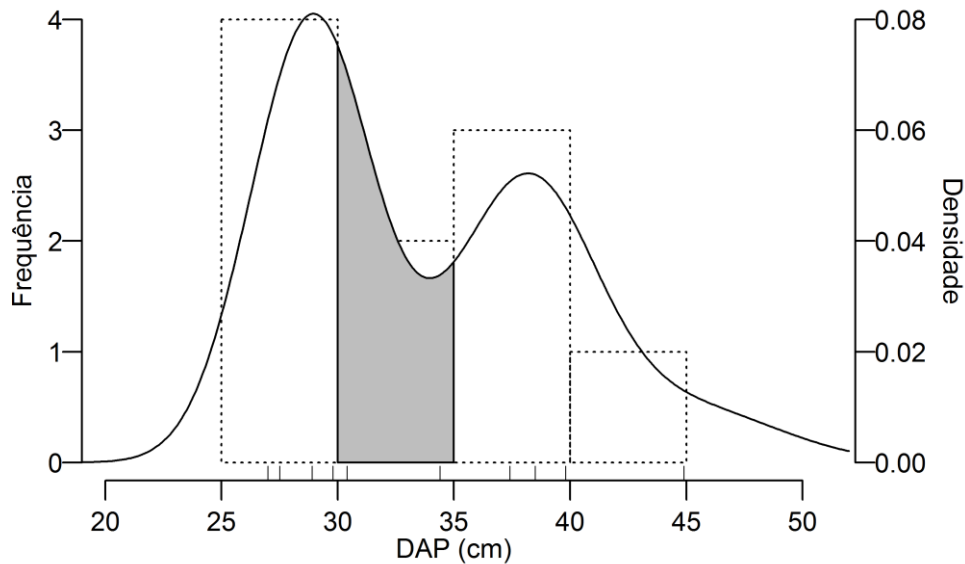


FIGURA 56 – PROBABILIDADE ENTRE 30 E 35 CM.  
FONTE: O AUTOR (2014)

Para exemplificar a geração de números pseudoaleatórios com base no método núcleo estimador variável, utilizaram-se os mesmos números sorteados de uma distribuição uniforme, conforme apresentado na Figura 19. O primeiro número sorteado foi 0,2540337. A fim de se calcular a inversa dessa distribuição para 0,2540337, considera-se  $\hat{F}(x) = P(X \leq x) = 0,2540337$  na equação (14):

$$\begin{aligned}
\hat{F}(x) &= 0,2540337 \therefore \\
0,2540337 &= \frac{1}{10} \sum_{i=1}^{10} F_{\mu=X_i; \sigma=2,455 \cdot \lambda_i} \therefore \\
0,2540337 &= \frac{1}{10} \cdot \left( F_{\mu=27,0; \sigma=2,455 \cdot 0,92}(x) + F_{\mu=27,5; \sigma=2,455 \cdot 0,88}(x) + F_{\mu=28,9; \sigma=2,455 \cdot 0,83}(x) \right. \\
&\quad + F_{\mu=29,8; \sigma=2,455 \cdot 0,85}(x) + F_{\mu=30,4; \sigma=2,455 \cdot 0,88}(x) + F_{\mu=34,4; \sigma=2,455 \cdot 1,14}(x) \\
&\quad + F_{\mu=37,4; \sigma=2,455 \cdot 1,01}(x) + F_{\mu=38,5; \sigma=2,455 \cdot 1,00}(x) + F_{\mu=39,8; \sigma=2,455 \cdot 1,07}(x) \\
&\quad \left. + F_{\mu=44,9; \sigma=2,455 \cdot 1,63}(x) \right)
\end{aligned} \tag{37}$$

A equação (37) deve ser resolvida por um algoritmo numérico. Por meio da função “optim” presente no software R (R CORE TEAM, 2014) e que implementa o algoritmo de Brent (1973), pode-se obter o valor de  $x$  na equação acima. Nesse caso, o valor obtido foi de 28,77. A Tabela 4 ilustra a obtenção de dez números pseudoaleatórios com base na função densidade de probabilidade obtida nesse exemplo por meio do método da inversa da cumulativa.

TABELA 4 – EXEMPLO DE GERAÇÃO DE NÚMEROS PSEUDOALEATÓRIOS COM BASE NO MÉTODO NÚCLEO ESTIMADOR VARIÁVEL  
FONTE: O AUTOR (2014)

$P(X \leq x)$	$x$
0,25	28,77
0,64	36,32
0,96	45,76
0,55	34,08
0,98	48,73
0,51	32,88
0,93	43,93
0,42	31,13
0,49	32,25
0,38	30,41

### APÊNDICE 3 – DIAGRAMA ENTIDADE RELACIONAMENTO

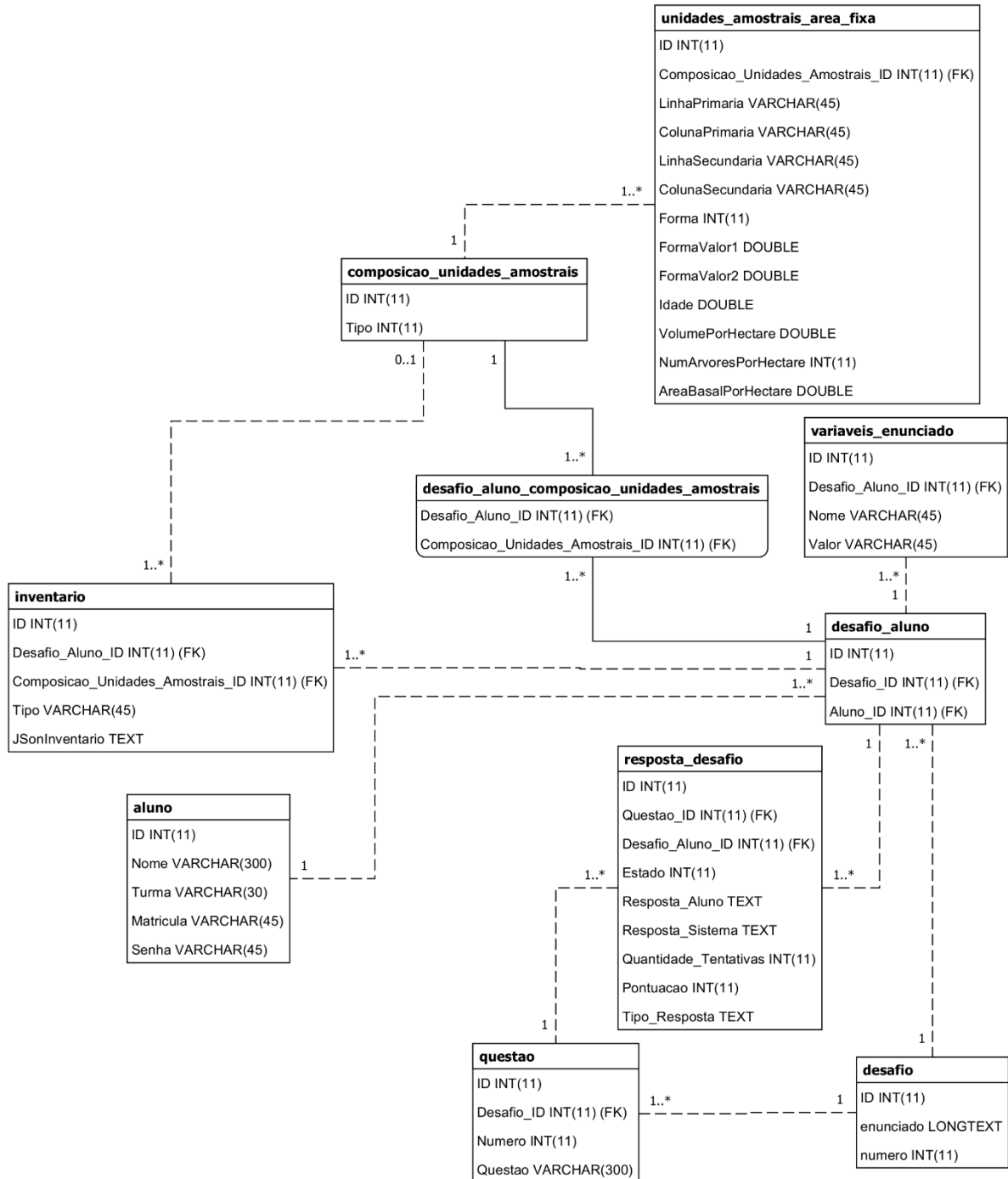


FIGURA 57 – DIAGRAMA DE ENTIDADE-RELACIONAMENTO – MODELO FÍSICO

## APÊNDICE 4 – QUADRO DAS CLASSES DA BIBLIOTECA DE PROGRAMAÇÃO PARA OS PROCESSOS DE AMOSTRAGEM

<b>Classe e Descrição</b>	<b>Revisão</b>	<b>Modelagem</b>
AmostragemEmConglomerados Classe que representa o processo de amostragem em conglomerados.	2.1.6	4.1.4
AmostragemEmDoisEstagios Classe que representa o processo de amostragem em dois estágios.	2.1.5	4.1.4
AmostragemEstratificada Classe que representa o processo de amostragem estratificada simples.	2.1.3	4.1.4
AmostragemEstratificadaAbstrata<T extends AmostragemSimples> Classe genérica que representa uma amostragem estratificada.	2.1.3	4.1.4
AmostragemEstratificadaTamanhosDiferentes Classe que representa o processo de amostragem estratificado com tamanhos diferentes.	2.1.3	4.1.4
AmostragemMOComRepeticaoParcial Classe que representa a amostragem em múltiplas ocasiões com repetição parcial.	2.1.8	4.1.4
AmostragemMODupla Classe que representa a amostragem em múltiplas ocasiões dupla.	2.1.8	4.1.4
AmostragemMOIndependentes Classe que representa a amostragem em múltiplas ocasiões independentes.	2.1.8	4.1.4
AmostragemMOREpeticaoTotal Classe que representa a amostragem em múltiplas ocasiões com repetição total.	2.1.8	4.1.4
AmostragemMultiplasOcasioes Classe que representa a generalização da amostragem em múltiplas ocasiões.	2.1.8	4.1.4
AmostragemSimples Essa classe é utilizada para os cálculos das estimativas de um conjunto de unidades amostrais da população que tenham igual chance de participar da amostra, caracterizando, assim, o processo de amostragem aleatória simples.	2.1.2	4.1.4
AmostragemSistematicaComMultiplosIniciosAleatorios Classe que representa o processo de amostragem sistemático com múltiplos inícios aleatórios.	2.1.4	4.1.4
AmostragemSistematicaEmDoisEstagios Classe que representa o processo de amostragem sistemático em dois estágios.	2.1.4	4.1.4
AmostragemSistematicaEmEstagioUnicoPorFaixa Classe que representa o processo de amostragem sistemático em estágio único por faixa.	2.1.4	4.1.4
AmostragemSistematicaEmEstagioUnicoPorParcela Classe que representa o processo de amostragem sistemática em estágio único e por parcela.	2.1.4	4.1.4
EstratoPorFaixa Classe que representa um conjunto de unidades amostrais que estão localizadas na mesma faixa geográfica.	2.1.3	4.1.4
EstratoPorIdade Classe que representa um conjunto de unidades amostrais que possuem árvores da mesma idade.	2.1.3	4.1.4
EstratoPorUnidadePrimaria Classe que representa um conjunto de unidades amostrais pertencentes à uma unidade amostral primária.	2.1.3	4.1.4

<b>Classe e Descrição</b>	<b>Revisão</b>	<b>Modelagem</b>
<b>EstratoSimples</b> Classe que representa um estrato de unidades amostrais de área fixa uma população ou amostra.	2.1.3	4.1.4
<b>Populacao</b> Esta classe representa uma população de unidades amostrais, onde as mesmas estão organizadas em linhas e colunas.	2.1.1	4.1.4
<b>PopulacaoPorFaixa</b> Classe que calcula os parâmetros da população estratificada por faixas.	2.1.1	4.1.4
<b>PopulacaoPorIdade</b> Classe que calcula os parâmetros da população estratificados por idade.	2.1.1	4.1.4
<b>PopulacaoPorUnidadePrimaria</b> Classe que calcula os parâmetros da população estratificados por unidade primária.	2.1.1	4.1.4
<b>ComposicaoUnidadesAmostrais&lt;T extends UnidadeAmostrais&gt;</b> Classe que representa uma composição de unidades amostrais, que também pode ser chamada de amostra.	-	4.1.3
<b>ComposicaoUnidadesAmostraisVO</b> Classe que representa uma amostra simples de unidades amostrais de área fixa.	-	4.1.3
<b>ConversorIndices</b> Classe utilitária que converte os índices das unidades amostrais.	-	-
<b>Forma</b> Classe que representa a forma geométrica de uma unidade amostral.	-	-
<b>Retangulo</b> Classe que representa a forma de um retângulo e é utilizada na unidade amostral de área fixa.	-	-
<b>UnAmostraiLinhaColunaComparador</b> Classe que implementa a interface Comparator utilizada para ordenar as UnidadeAmostraiAreaFixa.	-	-
<b>UnidadeAmostrai</b> Classe que representa uma unidade amostral.	-	4.1.1
<b>UnidadeAmostraiAreaFixa</b> Classe que representa uma unidade amostral de área fixa.	-	4.1.1

QUADRO 29 – CLASSES IMPLEMENTADAS NA BIBLIOTECA DE PROGRAMAÇÃO COM UMA BREVE DESCRIÇÃO  
FONTE: O AUTOR (2014)

O Quadro 29 apresenta as classes implementadas na biblioteca de programação para os processos de amostragem. Além do nome e descrição da classe o quadro apresenta mais duas colunas: “Revisão” que se refere à seção na qual o processo de amostragem foi revisado neste trabalho e “Modelagem” que se refere à seção na qual a modelagem da classe foi apresentada e discutida. Em algumas classes, essas colunas não apresentam valor, pois, as classes não fazem parte de um processo de amostragem específico, ou são classes utilitárias e simples.

## APÊNDICE 5 – EVIDÊNCIA DOS TESTES REALIZADOS NA BIBLIOTECA DE PROGRAMAÇÃO PARA OS PROCESSOS DE AMOSTRAGEM

O Quadro 30 apresenta parte dos valores obtidos nos testes da biblioteca de programação quando executada para resolver os exercícios propostos no livro texto de Inventário Florestal de Péllico Netto e Brena (1997, p. 97-300). Os resultados obtidos pela biblioteca de programação foram comparados com os resultados dos exercícios resolvidos no livro texto citado.

-----Estatísticas de toda a população-----						
Média =22,550						
Variância = 65,630						
Desvio Parão = 8,101						
Coeficiente de Variação = 35,926						
Volume Total = 10.147,500						
-----Estatísticas da população estratificada por idade-----						
Estrato(h)	NumUnidades(Nh)	Média(Xh)	Variância(S2h)	Desvio Padrão(Sh)	Coef de Variação(CVh)	
EI_6.0	144	13,134	19,508	4,417	33,629	
EI_9.0	164	25,241	21,445	4,631	18,346	
EI_12.0	142	28,990	23,673	4,865	16,783	
Média estratificada =0,000						
Variância estratificada= 0,000						
Variância entre os estratos = 553,832						
Variância dentro dos estratos = 21,528						
Volumes totais por estrato						
EI_6.0=1.891,300						
EI_9.0=4.139,600						
EI_12.0=4.116,600						
-----Estatísticas da população estratificada por faixa igual-----						
No.	F Col	Média por subuni(m3/0,1ha)	Variância (m3/0,1ha)^2	Volume por f (m3)		
1	a	23,283	68,065	698,500		
2	b	22,397	81,596	671,900		
3	c	22,807	75,211	684,200		
4	d	22,530	82,521	675,900		
5	e	22,143	59,987	664,300		
6	f	22,067	69,747	662,000		
7	g	21,857	62,498	655,700		
8	h	21,953	60,163	658,600		
9	i	22,097	79,416	662,900		
10	j	23,423	67,892	702,700		
11	k	22,837	64,849	685,100		
12	l	22,227	56,756	666,800		
13	m	21,877	61,517	656,300		
14	n	23,543	61,656	706,300		
15	o	23,210	59,194	696,300		
Média =676,500						
Variância = 314,609						
Desvio Padrão das faixas = 17,737						
Coef. de Variação = 2,622						
-----Estatísticas da população por unidade amostral primária-----						
Unidade Primária	Média (m3/0,1ha)	Variância(m3/0,1ha)^2	Coeficiente de Variação			
EUP_I_A	9,652	3,363	19,001			
EUP_I_B	16,196	15,264	24,123			
EUP_I_C	23,724	17,181	17,472			
EUP_I_D	26,976	17,726	15,607			
EUP_I_E	28,064	24,077	17,485			
EUP_I_F	31,180	20,857	14,647			
EUP_II A	10,160	10,160	31,373			

EUP_II_B	16,200	20,429	27,900
EUP_II_C	22,480	9,961	14,039
EUP_II_D	27,000	20,873	16,921
EUP_II_E	27,768	19,739	16,000
EUP_II_F	30,068	22,095	15,633
EUP_III_A	11,388	6,994	23,222
EUP_III_B	17,344	20,963	26,399
EUP_III_C	24,128	18,376	17,767
EUP_III_D	26,736	26,325	19,191
EUP_III_E	26,608	19,457	16,578
EUP_III_F	30,228	20,340	14,920
Variância Dentro das Unidades Primárias=17,455			
Variância Entre as Unidades Primárias=589,310			

QUADRO 30 – EVIDÊNCIA DOS TESTES REALIZADOS NA BIBLIOTECA DE PROGRAMAÇÃO QUANDO EXECUTADA PARA RESOLVER OS EXERCÍCIOS PROPOSTOS NO LIVRO TEXTO DE INVENTÁRIO FLORESTAL DE PÉLICO NETTO E BRENA (1997, P. 97-300)

FONTE: O AUTOR (2014)

A Figura 58 e a Figura 59 apresentam o resultado dos testes unitários programados no ambiente de desenvolvimento. A Figura 58 apresenta a tela que reporta um caso em que houve 4 erros e 3 falhas nos testes, o que pode ser visto nos itens em vermelho e em azul. Pode-se perceber que foram realizados 16 testes (que se referem aos 16 processos de amostragem implementados). As 3 falhas se referem aos valores que foram calculados e não estão de acordo com o que era esperado. Os 4 erros se referem a problemas internos da biblioteca de programação ou dos testes. Durante o desenvolvimento, quando ocorriam essas situações o software era corrigido para que voltasse a funcionar normalmente.

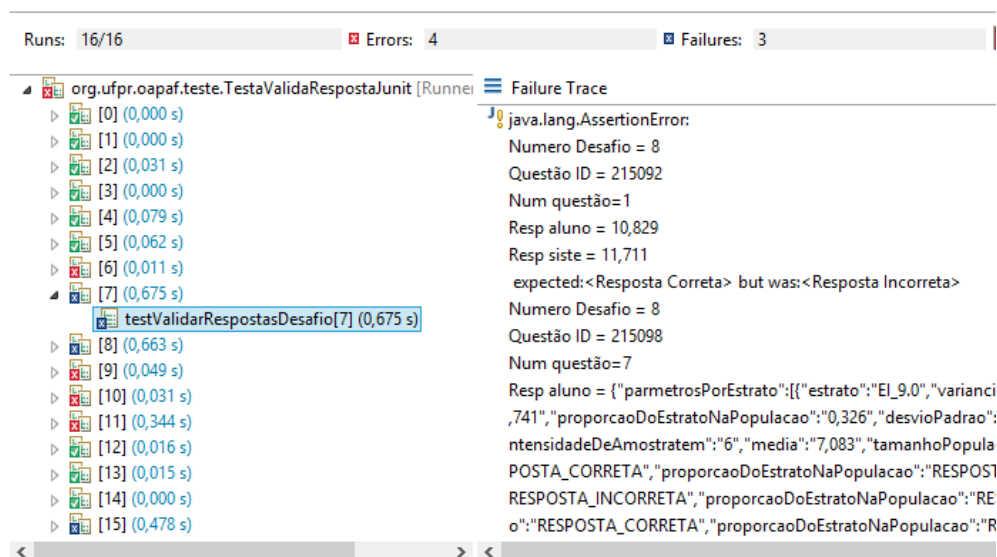


FIGURA 58 – EVIDÊNCIA DOS TESTES UNITÁRIOS COM ERROS E FALHAS  
FONTE: O AUTOR (2014)



A Figura 59 apresenta um caso em que os testes foram executados sem erros ou falhas. Esses testes foram executados para validar a biblioteca de programação para dados diferentes, bem como, para validá-la durante o desenvolvimento e garantir que novas funcionalidades não afetassem o que já estava funcionando.

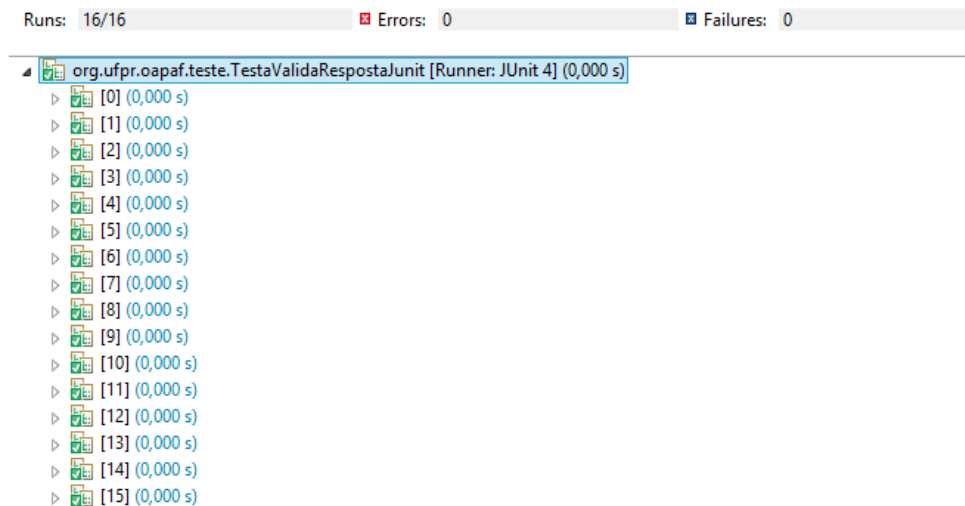


FIGURA 59 – EVIDÊNCIA DOS TESTES UNITÁRIOS SEM ERROS OU FALHAS  
FONTE: O AUTOR (2014)

**APÊNDICE 6 – EXEMPLO DE DADOS GERADOS PARA O EXERCÍCIO 1 DO  
“ALUNO 1”**

		I						II						III					
		a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r
A	1	218	145	183	94	109	151	88	138	172	151	132	75	142	97	83	108	100	154
	2	100	187	151	183	96	239	72	108	189	137	107	151	126	159	82	166	182	82
	3	96	123	97	115	62	180	112	105	134	104	124	213	87	154	153	127	132	72
	4	70	109	72	120	84	79	110	127	95	78	193	135	142	185	96	153	162	115
	5	139	146	103	155	166	84	67	83	100	102	187	155	62	217	142	68	147	95
B	6	130	81	119	97	81	101	142	112	177	233	83	86	235	156	147	161	213	162
	7	112	126	144	153	183	89	240	103	108	99	127	85	86	157	157	123	156	65
	8	87	83	121	108	166	107	119	109	85	108	115	123	123	153	87	232	81	86
	9	202	122	108	79	89	96	129	161	187	130	99	195	234	121	123	98	115	193
	10	82	135	199	168	96	88	72	133	75	121	159	167	166	85	96	105	226	126
C	11	169	96	65	121	76	137	181	69	96	119	161	238	182	133	229	87	87	196
	12	95	90	119	194	153	101	91	84	154	109	118	73	192	193	124	95	74	67
	13	124	189	99	126	191	230	332	254	307	230	231	275	198	263	288	275	258	183
	14	267	221	204	229	225	278	237	312	261	307	221	363	180	230	298	263	229	275
	15	251	271	190	294	292	190	220	247	221	238	168	290	235	230	260	229	249	337
D	16	193	320	261	235	265	323	173	267	177	195	198	209	247	264	236	211	319	273
	17	265	262	234	289	283	282	261	189	233	300	281	280	181	284	224	230	291	192
	18	184	247	270	208	264	228	255	197	234	238	175	259	260	207	295	277	195	323
	19	360	204	275	169	333	217	347	250	176	221	250	207	206	241	249	183	312	214
	20	277	304	274	291	242	309	266	268	204	293	218	325	187	281	245	310	244	250
E	21	274	301	288	248	236	323	171	262	201	327	233	268	194	262	188	202	293	278
	22	230	249	292	170	270	258	248	193	207	228	258	232	166	308	219	193	250	358
	23	337	228	194	196	290	335	252	277	196	347	233	262	312	280	203	177	225	278
	24	249	193	257	215	269	231	212	181	210	221	275	256	210	281	273	236	196	223
	25	386	320	277	237	250	174	378	309	346	254	268	315	232	303	327	315	297	215
F	26	307	257	226	266	247	317	260	350	300	347	256	415	213	267	342	302	265	315
	27	287	328	225	336	334	224	256	269	257	276	201	331	272	252	300	264	286	367
	28	227	361	300	272	305	365	206	321	210	229	232	246	269	304	270	248	360	314
	29	305	316	257	331	323	338	300	223	256	341	337	320	213	323	245	265	333	226
	30	217	285	311	246	304	250	290	232	271	276	226	293	300	245	338	334	229	365
G	31	408	242	315	221	362	253	377	286	227	243	273	245	244	279	286	215	350	251
	32	316	344	314	334	281	348	320	309	242	336	254	368	221	320	283	368	283	287
	33	331	342	327	270	273	365	204	316	235	370	270	308	228	302	222	241	335	317
	34	266	286	334	216	311	297	270	228	229	264	312	254	199	347	254	228	273	403
	35	382	263	228	230	332	380	274	319	236	399	270	301	351	320	225	210	247	317

QUADRO 31 – EXEMPLO DE DADOS GERADOS PARA O EXERCÍCIO 1 DO “ALUNO 1”  
FONTE: O AUTOR (2014)

O Quadro 31 apresenta um exemplo de dados gerados pelo método apresentado na seção 3.3 e como complemento aos resultados apresentados na seção 4.2.

Os dados gerados apresentados no Quadro 31 representa uma população de *Pinus sp* com 47,25 ha de área total, enumerada totalmente, por meio da divisão de 630 unidades amostrais de forma retangular, com 30 metros de largura por 25 metros de comprimento, ou seja 750 m<sup>2</sup> (0,075 ha) de superfície. Em cada uma das unidades amostrais são representados os volumes com casca, por hectare. A população foi estruturada em linhas numeradas de 1 até 35 e colunas, identificadas pelas letras “a” até “r”. Cada combinação entre as linhas (“1” até “35”) e colunas (“a” até “r”) representa uma unidade amostral secundária. A mesma população foi dividida em 21 unidades primárias, identificadas pela combinação dos algarismos romanos “I” até “III” e pelas letras maiúsculas “A” até “G”. Cada unidade primária contém 30 unidades secundárias. Além disso esse povoamento possui árvores em diferentes idades. As idades podem ser identificadas pelas cores verde claro (6 anos), verde (9 anos) e verde escuro (12 anos).

## APÊNDICE 7 – TESTE DE KOLMOGOROV-SMIRNOV PARA 102 CONJUNTOS DE DADOS GERADOS

TABELA 5 – RESULTADO DO TESTE DE KOLMOGOROV-SMIRNOV PARA 102 CONJUNTOS DE DADOS GERADOS  
FONTE: O AUTOR(2014)

Aluno	Idade	Dtab	Dcal	Aluno	Idade	Dtab	Dcal
Aluno 1	6 anos	0,15	0,03	Aluno 18	6 anos	0,16	0,04
Aluno 1	9 anos	0,14	0,03	Aluno 18	9 anos	0,15	0,06
Aluno 1	12 anos	0,15	0,03	Aluno 18	12 anos	0,15	0,06
Aluno 2	6 anos	0,16	0,08	Aluno 19	6 anos	0,16	0,06
Aluno 2	9 anos	0,15	0,04	Aluno 19	9 anos	0,15	0,05
Aluno 2	12 anos	0,15	0,03	Aluno 19	12 anos	0,15	0,02
Aluno 3	6 anos	0,16	0,08	Aluno 20	6 anos	0,15	0,04
Aluno 3	9 anos	0,15	0,07	Aluno 20	9 anos	0,15	0,02
Aluno 3	12 anos	0,15	0,04	Aluno 20	12 anos	0,15	0,02
Aluno 4	6 anos	0,15	0,03	Aluno 21	6 anos	0,15	0,05
Aluno 4	9 anos	0,15	0,05	Aluno 21	9 anos	0,15	0,04
Aluno 4	12 anos	0,15	0,07	Aluno 21	12 anos	0,15	0,02
Aluno 5	6 anos	0,16	0,02	Aluno 22	6 anos	0,16	0,03
Aluno 5	9 anos	0,15	0,03	Aluno 22	9 anos	0,15	0,04
Aluno 5	12 anos	0,15	0,02	Aluno 22	12 anos	0,15	0,04
Aluno 6	6 anos	0,15	0,03	Aluno 23	6 anos	0,15	0,07
Aluno 6	9 anos	0,14	0,10	Aluno 23	9 anos	0,14	0,05
Aluno 6	12 anos	0,16	0,06	Aluno 23	12 anos	0,15	0,06
Aluno 7	6 anos	0,15	0,03	Aluno 24	6 anos	0,15	0,07
Aluno 7	9 anos	0,14	0,03	Aluno 24	9 anos	0,15	0,07
Aluno 7	12 anos	0,15	0,03	Aluno 24	12 anos	0,16	0,02
Aluno 8	6 anos	0,15	0,08	Aluno 25	6 anos	0,15	0,09
Aluno 8	9 anos	0,16	0,09	Aluno 25	9 anos	0,15	0,06
Aluno 8	12 anos	0,15	0,05	Aluno 25	12 anos	0,15	0,03
Aluno 9	6 anos	0,16	0,03	Aluno 26	6 anos	0,15	0,09
Aluno 9	9 anos	0,16	0,02	Aluno 26	9 anos	0,15	0,05
Aluno 9	12 anos	0,16	0,01	Aluno 26	12 anos	0,15	0,06
Aluno 10	6 anos	0,15	0,04	Aluno 27	6 anos	0,15	0,02
Aluno 10	9 anos	0,15	0,07	Aluno 27	9 anos	0,15	0,04
Aluno 10	12 anos	0,15	0,04	Aluno 27	12 anos	0,15	0,03
Aluno 11	6 anos	0,15	0,03	Aluno 28	6 anos	0,15	0,03
Aluno 11	9 anos	0,15	0,07	Aluno 28	9 anos	0,14	0,05
Aluno 11	12 anos	0,15	0,03	Aluno 28	12 anos	0,15	0,05
Aluno 12	6 anos	0,15	0,02	Aluno 29	6 anos	0,15	0,06
Aluno 12	9 anos	0,15	0,03	Aluno 29	9 anos	0,15	0,05
Aluno 12	12 anos	0,16	0,02	Aluno 29	12 anos	0,15	0,03
Aluno 13	6 anos	0,15	0,04	Aluno 30	6 anos	0,15	0,03
Aluno 13	9 anos	0,15	0,03	Aluno 30	9 anos	0,14	0,05
Aluno 13	12 anos	0,16	0,03	Aluno 30	12 anos	0,15	0,03
Aluno 14	6 anos	0,15	0,04	Aluno 31	6 anos	0,15	0,06
Aluno 14	9 anos	0,14	0,02	Aluno 31	9 anos	0,15	0,09
Aluno 14	12 anos	0,15	0,05	Aluno 31	12 anos	0,15	0,09
Aluno 15	6 anos	0,15	0,02	Aluno 32	6 anos	0,15	0,05
Aluno 15	9 anos	0,15	0,03	Aluno 32	9 anos	0,15	0,05
Aluno 15	12 anos	0,15	0,02	Aluno 32	12 anos	0,15	0,04
Aluno 16	6 anos	0,16	0,04	Aluno 33	6 anos	0,15	0,03
Aluno 16	9 anos	0,15	0,04	Aluno 33	9 anos	0,15	0,03
Aluno 16	12 anos	0,16	0,04	Aluno 33	12 anos	0,14	0,06
Aluno 17	6 anos	0,15	0,04	Aluno 34	6 anos	0,16	0,08
Aluno 17	9 anos	0,15	0,04	Aluno 34	9 anos	0,15	0,03
Aluno 17	12 anos	0,15	0,03	Aluno 34	12 anos	0,16	0,03

## APÊNDICE 8 – RESPOSTA DOS ALUNOS AOS ITENS ABERTOS DO INSTRUMENTO DE AVALIAÇÃO

Item 11 - O que você gostou no Simulador Florestal?	
Aluno	Resposta
1	"O sistema de pontuação de acordo com o número de acertos."
4	"Fácil utilização e linguagem."
5	"Software de fácil utilização, acrescentando prática a disciplina."
6	"Facilita o entendimento da matéria."
7	"Gostei de tudo."
8	"A apresentação de formulário para auxílio nos testes, correção das questões logo em seguida para confirmação do valor."
9	"Interação imediata"
10	"Gostei de ter dados diferentes de meus colegas e poder saber na hora em qualquer lugar se acertei ou não e também das pontuações e fórmulas."
11	"Da correção automática."
12	"A correção em tempo real e o formulário disponível para download."
13	"A facilidade do arquivo, facilitando o entendimento no Excel."
15	"A possibilidade de fixar o conteúdo da disciplina de maneira interativa."
16	"A disponibilização de dados realistas e possibilidade de conferir os resultados."
17	"A conexão em tempo real, que proporciona uma grande interatividade com o software."
18	"Correção Simultânea."
19	"Gostei da correção que o simulador proporciona, ou seja, uma correção imediata. Da lista de fórmulas que o simulador disponibiliza e a facilidade de utilização."
20	"A facilidade, tipo de pontuação, a necessidade de avaliar bem antes de colocar a resposta."
21	"A possibilidade de corrigir as respostas em tempo real."
22	"A correção em tempo real dos exercícios."
Item 12 - O que você não gostou no Simulador Florestal?	
Aluno	Resposta
1	"Na minha opinião está excelente."
5	"Sistema de pontuação."
8	"Ranking de pontuação."
9	"Excesso de cálculos e pouca adequação a realidade dos inventários."
10	"Nada."
11	"Interface."
12	"Precisa voltar à tela inicial para conferir dados do enunciado."
20	"Nada, o simulador é muito bom."
21	"De ter que exportar os arquivos para o excel."
Item 13 - Você tem alguma sugestão sobre o Simulador Florestal?	
Aluno	Resposta
6	"Não, pois está adequado para utilizar."
8	"Formulário para todos os testes."
10	"Colocar sons para erros e acertos, mas podendo desativá-los."
13	"Mais informações para pessoas com pouca prática no Excel."
19	"Possibilidade de inserção de dados."
20	"Poderia complementar com outras partes do inventário: ANATRO completa, por exemplo."
21	"Deixar a fórmula a ser utilizada ao lado do valor a ser respondido, assim podendo excluir o formulário."
22	"Seria mais interessante se fosse possível inserir dados no software."

QUADRO 32 – RESPOSTAS DOS ALUNOS NOS ITENS ABERTOS (11 A 13) DO INSTRUMENTO DE AVALIAÇÃO

FONTE: O AUTOR (2014)

**ANEXO 1 – QUADRO COM DADOS ORIGINAIS DO LIVRO TEXTO DE PÉLLICO  
NETTO E BRENA (1997)**

		I					II					III				
		a	b	c	d	e	f	g	H	i	j	k	l	m	n	o
A	1	80	92	96	94	90	85	73	63	83	101	115	156	87	109	111
	2	99	69	102	103	91	123	83	128	68	98	86	88	95	97	74
	3	86	69	85	127	98	102	98	179	71	116	98	101	88	125	110
	4	81	89	122	110	80	99	184	81	85	114	191	132	122	110	156
	5	131	115	92	76	136	157	95	80	89	85	126	106	104	144	116
B	6	162	100	118	90	116	83	163	95	107	125	145	162	87	225	255
	7	166	164	191	190	165	155	186	188	156	108	116	177	229	149	127
	8	185	227	171	239	185	114	138	186	232	213	147	125	159	170	197
	9	216	101	148	151	149	159	158	184	142	180	159	126	162	199	156
	10	189	197	132	137	160	190	165	240	125	258	205	214	204	157	284
C	11	236	269	172	237	243	213	233	205	244	230	229	238	240	310	284
	12	273	176	217	194	314	221	201	193	239	184	162	173	216	211	254
	13	197	279	225	184	237	169	228	204	253	271	210	232	195	322	209
	14	246	256	249	180	231	229	188	199	200	242	221	274	307	272	191
	15	306	281	248	294	187	196	278	241	272	287	263	229	305	241	244
D	16	267	223	284	213	239	235	203	246	307	264	236	199	227	219	176
	17	204	256	273	246	279	259	192	221	294	282	291	232	199	259	256
	18	253	228	259	263	292	239	223	335	359	259	319	244	307	351	295
	19	280	256	292	386	289	327	283	219	232	349	326	262	229	253	331
	20	324	273	365	268	232	266	249	317	298	292	246	358	226	305	338
E	21	301	268	323	276	289	347	231	278	205	284	213	243	214	339	296
	22	402	241	360	399	278	346	247	279	253	366	248	335	283	249	229
	23	226	255	229	247	269	242	267	207	233	317	336	225	287	207	229
	24	305	255	257	210	265	270	337	307	318	228	314	321	224	297	238
	25	267	239	298	248	309	279	269	253	261	318	271	322	218	234	280
F	26	318	306	327	320	255	258	242	228	266	292	309	263	262	379	322
	27	318	329	248	287	267	273	339	345	272	283	348	221	307	262	280
	28	292	415	287	259	255	266	384	336	363	311	267	313	330	232	235
	29	255	314	335	331	273	339	351	325	257	301	286	285	283	278	342
	30	320	377	337	400	370	379	269	224	345	269	368	312	367	358	348

QUADRO 33 – DADOS EM M<sup>3</sup>/HA DE UMA POPULAÇÃO DE *PINUS SP*  
FONTE: PÉLLICO NETTO E BRENA (1997, P. 98)

O Quadro 33 representa uma população de *Pinus sp* com 45 ha de área total, enumerada totalmente, por meio da divisão de 450 unidades amostrais de forma retangular, com 20 metros de largura por 50 metros de comprimento, ou seja 1.000m<sup>2</sup> (0,1 ha) de superfície. Em cada uma das unidades amostrais são representados os volumes com casca, por hectare. A população foi estruturada em linhas numeradas

de 1 até 30 e colunas, identificadas pelas letras “a” até “o”. Cada combinação entre as linhas (“1” até “30”) e colunas (“a” até “o”) representa uma unidade amostral secundária. A mesma população foi dividida em 18 unidades primárias, identificadas pela combinação dos algarismos romanos “I” até “III” e pelas letras maiúsculas “A” até “F”. Cada unidade primária contém 25 unidades secundárias. Além disso esse povoamento possui árvores em diferentes idades. As idades podem ser identificadas pelas cores verde claro (6 anos), verde (9 anos) e verde escuro (12 anos) (PÉLLICO NETTO e BRENA, 1997, p. 98).